

## DOCUMENT RESUME

ED 445 084

TM 031 761

AUTHOR Seaman, Alan; Hayward, Linda  
TITLE Standardized ESL Test Equating Study: Equating the CELSA and the NYSPT with the MELT SPLs.  
SPONS AGENCY Department of Education, Washington, DC.; Illinois State Board of Education, Springfield.  
PUB DATE 2000-06-00  
NOTE 83p.; "Submitted by a Coalition representing the West Suburban Area Planning Council."  
PUB TYPE Numerical/Quantitative Data (110) -- Reports - Research (143)  
EDRS PRICE MF01/PC04 Plus Postage.  
DESCRIPTORS \*Academic Achievement; Adult Education; \*Adults; \*English (Second Language); \*Equated Scores; Limited English Speaking; \*Test Use

## ABSTRACT

The purpose of this study was to equate two standardized English as a Second Language (ESL) proficiency tests, the New York State Placement Test (NYSPT) and the Combined English Language Skills Assessment (CELSA), with the Mainstream English Language Training (MELT) project Student Performance Levels (SPLs) in order to facilitate the reporting of student progress in adult ESL programs. A random sample of 593 adult ESL students was drawn from 3 programs in the western suburbs of Chicago, Illinois, that are representative of the ESL programs across the state. Between November 1998 and May 1999, the students in this sample were tested in a counterbalanced random groups design, pairing the NYSPT with the Basic English Skills Test (BEST) Oral test and the CELSA with the BEST Literacy test. In a statistical analysis, highly positive correlations were found between the two oral tests (0.85) and between the two literacy tests (0.78). Since the two sections of the BEST had already been equated with the MELT SPLs, they were used to establish SPLs for each student in the sample. These data were then used to establish score ranges for the NYSPT and the CELSA at each SPL. In addition to equating these tests statistically with the SPLs, the research team collected qualitative data in the form of surveys and interviews in order to make recommendations for how the tests might be used to report student progress in adult ESL programs. Twelve appendixes contain statistical details and the forms used in the study. (Author/SLD)

## SCOPE OF INTEREST NOTICE

The ERIC Facility has assigned this document for processing to:

In our judgment, this document is also of interest to the Clearinghouses noted to the right. Indexing should reflect their special points of view.

TM  
FL

## Standardized ESL Test Equating Study: Equating the CELSA and the NYSPT with the MELT SPLs

### Authors:

**Alan Seaman, Ph.D. *Research Director***  
**Linda Hayward, *Project Director***

### Administrative Team:

**Joanna Escobar, *College of DuPage***  
**Marilyn Sweeny, *World Relief DuPage***  
**Don Zabelin, *West Chicago Community H.S. District 94***

U.S. DEPARTMENT OF EDUCATION  
Office of Educational Research and Improvement  
EDUCATIONAL RESOURCES INFORMATION  
CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

**Submitted by a Coalition representing the  
West Suburban Area Planning Council**

**College of DuPage  
West Chicago Community High School District 94  
World Relief DuPage**

**June 1999  
Revised, June 2000**

This research project was funded under the provisions of Section 353 of the Federal Education Act, and supported in whole by the U.S. Department of Education and the Illinois State Board of Education, Community and Family Partnerships Division. The facts and opinions stated in this report are those of the project staff and do not necessarily represent those of the Illinois State Board of Education.

**BEST COPY AVAILABLE**

## TABLE OF CONTENTS

### Abstract

### I. Introduction.

A. Why the Study Was Needed .....	2
B. The Tests .....	5
C. The Agencies .....	6

### II. Methodology.

A. Research Contexts and Time Frame .....	7
B. Sampling and Data Collection Procedures .....	8
C. Training and Test Administration .....	9
D. Statistical Analysis Procedures .....	10
E. Qualitative Analysis Procedures .....	11
F. Limitations of the Study .....	11

### III. Results and Discussion.

A. Supporting Data .....	12
B. Demographic Profiles of the Agencies .....	14

### IV. Recommendations for Use of the Tests. .... 16

### V. Conclusion. .... 17

### References

### APPENDICES.

Appendix A	Statistical Results for Equating
Appendix B	Statistical Profiles: Program Demographics
Appendix C	Comparison of Score Ranges for Three Different Proficiency Scales
Appendix D	Training Schedule and Materials
Appendix E	Testing Team Instructions
Appendix F	Testing Schedule
Appendix G	Letters to Teachers and Students
Appendix H	Consent Forms
Appendix I	Individual Student Biodata/Test Outcomes Forms
Appendix J	SPL Checklist for Phase 3
Appendix K	Team Members
Appendix L	Recommendations for Testing in Various Programs

## Abstract

The purpose of this study was to equate two standardized ESL proficiency tests, the NYS Place and the CELSA, with the MELT Student Performance Levels (SPLs), in order to facilitate the reporting of student progress in adult ESL programs. A random sample of 593 adult ESL students was drawn from three programs in the western suburbs of Chicago: the College of DuPage, World Relief DuPage, and West Chicago Community High School District 94. These programs are representative of the types of ESL programs that exist across Illinois. Between November 1998 and May 1999, the students in this sample were tested in a counterbalanced random groups design, pairing the NYS Place with the BEST Oral test and the CELSA with the BEST Literacy test. In a statistical analysis, high positive correlations were found between the two oral tests (.85) and between the two literacy tests (.78). Since the two sections of the BEST had already been equated with the MELT SPLs, they were used to establish SPLs for each of the students in the sample. These data were then used to establish score ranges for the NYS Place and the CELSA at each Student Performance Level.. In addition to statistically equating these tests with the SPLs, the research team collected qualitative data in the form of surveys and interviews in order to make recommendations for how the tests might be used to report student progress in adult ESL programs.

## I. Introduction.

### A. *Why the Study Was Needed.*

The U.S. Office of Education (USOE) requires the Illinois State Board of Education (ISBE) to report test results for all adult education students. The test results may be reported using either CASAS test scores or the Student Performance Levels (SPLs) based on the MELT project. Consequently, Illinois requires that local programs test all adult students above SPL 3 using the CELSA, and that students below SPL 3 be tested with the BEST Literacy or the NYSPT. However, only one of these tests, the BEST Literacy, has been equated with the MELT SPLs. This research study seeks to answer a fundamental question related to these requirements: How do test scores on the CELSA and NYSPT translate into the Student Performance Levels?<sup>1</sup>

To date, no study has equated the SPLs with the CELSA or NYSPT. In addition to enabling adult education programs in Illinois to satisfy USOE reporting mandates, equating these two tests to the SPLs will provide consistency in reporting test scores in programs across the state. Equating the CELSA and NYSPT with the SPLs will also offer local adult education programs and students descriptions of language proficiency, rather than simply stating a test score which is virtually without meaning for students and programs alike. The SPLs provide a short catalog of linguistic behaviors which allow programs to develop performance outcomes statements and to assist in curriculum planning.

The importance of this study was highlighted in a recent monograph, *Research Agenda for Adult ESL*, published by the Center for Applied Linguistics and the National Clearinghouse for ESL Literacy Education (1998). Surveying the current assessment situation, the authors note that a number of standardized tests exist,

---

<sup>1</sup>The MELT project refers to the Mainstream English Language Training project sponsored by the Office of Refugee Resettlement (ORR). The BEST refers to the Basic English Skills Test; the CELSA refers to the Combined English Language Skills Assessment; the NYSPT refers to the New York State Placement Test.

but “there is no clear understanding of how scores on these different tests compare” (p. 10). As a result, there is a lack of consistency in assessing and reporting student progress. By equating two standardized tests with a widely-used SPL scale, this study will address the problem of comparability. And, by interpreting the results of this study, we hope to make recommendations for the use of the CELSA, BEST Literacy, and NYSPT with various types of programs, students, and levels.

To equate the CELSA and NYSPT with the SPLs, score ranges on the BEST Core (Oral) and BEST Literacy were used as an intermediate step, since these scores had already been equated with the SPLs (*BEST Test Manual*, pp. 8, 12). Pairs of tests (the CELSA/BEST Literacy and the NYSPT/BEST Core) were given to a large random sample of adult ESL students in three Illinois agencies: the College of DuPage, World Relief DuPage, and West Chicago Community High School District 94. The tests were administered consistently and systematically to 593 students over a seven-month period (November 1998-May 1999) by a testing team under the supervision of the project director. Strong positive correlations between the tests (.78 for the CELSA/BEST Literacy; .85 for the NYSPT/BEST Core) made it possible to establish reliable score ranges for the SPLs.

The three agencies conducting this study represent three major types of service delivery responsible for conducting adult ESL programs in Illinois. The College of DuPage offers a large number of ESL classes on its main campus and at various off-campus sites throughout DuPage County, drawing a broad range of students. (In addition, COD offers academic-professional ESL classes which were used to equate the CELSA with the highest levels of the SPLs.) World Relief DuPage is a smaller community-based organization focusing on refugee resettlement and education. West Chicago Community High School District 94 provides adult ESL instruction within a school district for a more homogeneous, largely Hispanic community.<sup>2</sup> These agencies provide external validity, a real-world environment, and access to an aggregate ESL student population representing an authentic cross-section of the ESL population served by local ISBE-funded adult education programs in Illinois.

The results of this study can be seen in Table 1, which provides ranges of scores for levels of the SPLs as described in the original literature on the SPLs published by the MELT project in the 1980s and in the technical manual for the BEST.

In the late 1990s, the SPLs were reformulated by Allene Grognet at the Center for Applied Linguistics to more carefully distinguish between the oral and literacy (reading/writing) levels (Grognet, 1997). These changes are reflected in Table 2.

---

<sup>2</sup>For the sake of efficiency, West Chicago Community High School District 94 will be referred to as “West Chicago District 94.”

**Table 1: Score Ranges for Standardized ESL Tests**

<b>MELT SPL</b>	<b>BEST Literacy (Scaled Score)</b>	<b>CELSA</b>	<b>NYS Place</b>	<b>General Level</b>
0	0-2	---	--	Pre-Literacy
I	3-7	---	0-3	Literacy
II	8-21	---	4-6	Low Beginning
III	22-35	---	7-14	High Beginning
IV	36-46	20-23	15-20	Low Intermediate
V	47-53	24-29	21-27	Intermediate
VI	54-65	30-41	28-32	High Intermediate
VII	66+	42-53	33 +	Low Advanced
VIII	---	54-64	---	Advanced
IX	---	65+	---	High Advanced

**Table 2: Score Ranges in Terms of New Federal Levels for the SPLs**

<b>New Federal Levels</b>	<b>SPL Reading Writing</b>	<b>BEST Literacy</b>	<b>CELSA</b>	<b>SPL Speaking Listening</b>	<b>NYS Place</b>	<b>BEST Oral</b>
Beginning ESL Literacy	0-1	0-7	----	0-1	0-3	0-15
Beginning ESL	2-4	8-46	20-23	2-3	4-14	16-41
Low Intermediate ESL	5	47-53	24-29	4	15-20	42-50
High Intermediate ESL	6	54-65	30-41	5	21-27	51-57
Advanced	7	66+	42-53	6	28-32	58-64
Proficient	8+	----	54+	7+	33 +	65+

## ***B. The Tests.***

*The CELSA* is a 75-item multiple choice test which uses a cloze reading format. It was developed in the late 1980s by Donna Ilyin by combining the three 25-item tests which made up the ELSA. As a comprehensive English language skills test, the CELSA is intended to provide an efficient means of assessing students at the beginning, intermediate, and advanced levels, and it can be used for both placement and recording progress. The intended audience for this test is broad: adult resident immigrants, as well as students in community college, university, and high school ESL programs. The CELSA is given in a one-hour period, with 15 minutes for instructions and 45 minutes to take the test. The test answers may be recorded on scantron sheets.

As the CELSA user's guide notes, the test is not appropriate for pre-literate students at the low beginning level of proficiency. Since it primarily tests reading and grammar, the CELSA does not provide discrete point measurement of progress in life skills (e.g., reading a train schedule, telling time, using a map). According to the user's guide, the CELSA is "not suitable to show gains in any program that did not use materials that taught [reading and grammar in context]" (p. 3).

*The NYSPT*, also known as the NYS Place Test, is a 27-item test of oral English language proficiency to be used, as the title implies, primarily for the purpose of placement. It was developed in the 1980s by the New York State Education Department and became available for use in the early 1990s. The NYSPT is given as a one-to-one oral interview lasting approximately 10-15 minutes. Although the test has three sections -- the Oral Warm-up, the Basic English Literacy Screening, and the Oral Assessment with Pictures -- the third section is the main part of the test and the only section for which scores are reported for placement. The scoring of the test uses a simple 0-1-2 scale which is reliable but requires training. (The NYSPT packet includes a training videotape.) Like the CELSA, this test is intended to assess students across the beginning, intermediate, and advanced levels.

*The BEST* is a competency-based assessment of listening, speaking, reading, and writing skills which was developed in the late 1970s and early 80s by the Center for Applied Linguistics. It consists of two sections, the Core section (a 15-minute one-to-one interview) and the Literacy Skills section (a one-hour reading and writing test). Unlike the CELSA and NYSPT, the BEST focuses on specific functional tasks such as telling time, giving directions, reading a time table, and writing a note to a landlord. It was developed specifically for limited-English-speaking adults at the beginning and intermediate levels in programs which emphasize acquisition of English for life skills. In addition to providing scores for placement and progress assessment, the BEST can be used diagnostically to determine if students have mastered specific competencies covered in a life skills curriculum (*BEST Test Manual*, p. 1).

Because the BEST was field-tested by MELT project participants, both sections of the test were correlated with the Student Performance Levels. BEST score ranges for each SPL are provided in the Test Manual for the Core section (p. 8) and the Literacy Skills section (p. 12). Currently, only the BEST Literacy Skills section has been approved for reporting student progress in Illinois adult ESL programs. The BEST Literacy (Form B) is a 78-point test which is scored using both discrete point values and scales. Training is helpful for the BEST Literacy to be scored reliably.

For the purposes of this study, the two reading tests (CELSA and BEST Literacy) were paired, and the two oral tests (NYSPT and BEST Core) were paired. Although the content of the tests in each pair has somewhat different emphases, they are similar enough to expect the positive level of correlation necessary for equating.



### *C. The Agencies.*

As Peterson (1989) has noted, "in practice, whenever one refers to scores on two tests as being equated, it is important to add a qualifying phrase describing the population for which the equating is likely to hold" (p. 243). In other words, the characteristics of the population in this study are relevant to the generalizability of the results.

This study was conducted under the supervision of the West Suburban Area Planning Council, which is comprised of ESL administrators from the College of DuPage, West Chicago School District 94, and World Relief DuPage. These adult ESL programs represent the range of programs throughout the state. A more detailed analysis of the demographics of each program will be provided in Section 3 of this report.

*The College of DuPage* provides a district in DuPage, Cook, and Will Counties with an extensive ESL literacy program as well as academic-professional classes and an intensive English language institute. The literacy program includes 43 classes per term (four terms per year) in multiple sites throughout the district. The students in this program are diverse, representing a broad range of ethnic and educational backgrounds. In this study, literacy classes offered on the COD main campus were sampled, as well as classes in 4 of the 17 off-campus sites. Students in the highest-level classes of the academic-professional program also provided data for equating the CELSA with SPLs 8-9. The data collection period in this study covered parts of three terms.

*West Chicago Community High School District 94* offers adult ESL classes for a district which is 35 % Hispanic, consisting primarily of immigrants from Mexico. Thirteen classes are offered each semester at a high school in the evenings. The data collection period for this study covered two semesters, with samples drawn from a total of 22 classes.

The ESL services provided by *World Relief DuPage* are focused on refugees who are being resettled in the western suburbs. However, the classes are also open to immigrants on a space-available basis. World Relief DuPage offered 18 ESL classes at five sites in 1998-99, with students from four of these sites participating in the study. During the data collection period, a total of 29 classes were sampled for testing. Although the length of a term in this program is 10 months, because of a rolling admissions policy the classes tested in the spring were composed of different students than in the fall.

Table 3 provides a comparison of the three programs based on the ESL students sampled for this study. (It does not include the academic-professional classes at COD.)

In addition to the three agencies representing the West Suburban Planning Council, *Wheaton College* participated in the study in an advisory capacity through the research director, Dr. Alan Seaman. Several graduate students in the M.A. in Intercultural Studies/TESL program at Wheaton College participated as members of the testing team.



**Table 3: Demographic Comparison of the Sample from the Three Agencies**

Agency	Sample size (N)	Languages in sample	Countries in sample	Mean age of sample	Mean years of education	Gender of sample (percent)
West Chicago D94	124	4	4	28.17	9.43	M 62.9 F 37.1
World Relief	146	17	23	34.99	9.99	M 41.8 F 58.2
College of DuPage Literacy Program	323	29	39	32.43	10.88	M 42.7 F 57.3
Totals	593	37	49	32.17	10.35	M 46.7 F 53.3

## **II. Methodology.**

This equating study followed a *counterbalanced random groups design*, a common design for this type of study which controls for factors such as learning, fatigue, and practice.

### **A. Research contexts and time frame.**

The data collection period began on November 3, 1998 and ended on May 20, 1999. During this period, the tests were administered in three phases.

During *Phase 1 (November-December 1998)*, the testing team administered the CELSA and BEST Literacy to students who were sampled from intermediate-level classes in each program. Both tests were given to each student within a one-week time frame. The sequence of testing was counterbalanced, with approximately half of the students in the sample randomly assigned to take either the CELSA or the BEST Literacy first.

During *Phase 2 (January-May 1999)*, the testing team continued to administer the CELSA and BEST Literacy tests to randomly-sampled students from intermediate-level classes in each program. In addition, the CELSA and BEST Literacy were also given to advanced beginning-level classes across each program. During this phase, the testing team administered the NYSPT and BEST Core tests to a random sample of students across the beginning and low intermediate classes in each program. These tests were given back-to-back and were also counterbalanced. By the end of Phase 2, a total of 593 students had been tested, with 277 taking the BEST Literacy and CELSA, and 316 taking the BEST Core and NYSPT.

*Phase 3 (April-May 1999)*, which overlapped with Phase 2, focused on equating the CELSA with SPLs 8 and 9. ESL students in the highest level academic-professional classes at the College of DuPage. The testing team administered the CELSA to the entire population of students in these classes (n=103) and

their teachers concurrently rated each student on a checklist created from the descriptors in the MELT SPLs. During Phase 3, members of the testing team were debriefed about the tests used in this study through a questionnaire and a focus group interview which was recorded and analyzed.

### ***B. Sampling and Data Collection Procedures.***

The complete range of beginning and intermediate-level classes across the three agencies was used in this study. Each student on a class enrollment list was assigned a number and, using a set of random numbers, the research director sampled 42 % of the class. The sampled students were given information about the research study and a consent form; both the information and the consent form were translated into 11 languages. (See Appendices G and H). Those students who chose not to participate were not tested.

For the students sampled in Phase 1 and Phase 2, data were collected in the form of two test scores and a biodata sheet. The biodata sheet (see Appendix I) collected basic information about each student, including program, birthdate, sex, country of origin, first language, educational background, length of time in the US and length of English study. For the students in the academic-professional classes in Phase 3, data were collected in the form of one test score (the CELSA), a checklist based on the descriptors for SPLs 7-10 (filled out by the teacher) and the biodata sheet.

Every effort was made to collect the data in a highly systematic and consistent manner. The testing team underwent extensive training prior to each phase of data collection. Clearly-defined protocols were established for scheduling the testing, contacting teachers and students, sampling the class lists, administering the tests, completing the biodata forms, scoring the tests, and storing the data. Each part of this process was overseen by the project director, Linda Hayward, who was present at all testing sessions.

To minimize the effects of learning, both tests were given to each student within a short period of time: within one week for the CELSA and BEST Literacy, and within one hour for the NYSPT and the BEST Core. In addition, to minimize the effects of test-taking practice and fatigue, the tests were counterbalanced. This means that approximately half of the students sampled took one test first, and the other half of the sample took the tests in reverse order. The counterbalancing statistics are shown in Table 4.

**Table 4: Counterbalancing Data**

<b>Agency</b>	<b>BEST Literacy</b>	<b>CELSA</b>	<b>BEST Core</b>	<b>NYSPT</b>
West Chicago	18	28	39	37
D94	40	30	33	43
World Relief				
College of DuPage	81	78	83	81
<b>Totals</b>	<b>139</b>	<b>136</b>	<b>155</b>	<b>161</b>

### ***C. Training and Test Administration.***

In August and September 1998, a testing team was selected by the administrators of each ESL program, in consultation with the research director and project director. The team consisted of 13 experienced teachers from each of the three programs and graduate students in TESL from Wheaton College. The testing team was expanded to 16 in January 1999 in order to more efficiently administer the two one-to-one oral tests.

All of the members of the testing team received extensive training and practice prior to Phases 1 and 2 of data collection. This training was planned and conducted by the research director and the project director. In addition to familiarizing the team with the tests and establishing protocols for test administration, the training involved administering practice tests and recording the scoring of each team member. From these scores, inter-rater reliability was computed. It was decided that no data collection would take place until after an inter-rater reliability higher than .90 was established for the scorers of the two written tests, and an inter-rater reliability higher than .80 was established for all team members who scored the two oral tests. The inter-rater reliability figures are shown in Appendix A.

*Training for Phase 1:* Once the members of the testing team had been selected, training began with a two-hour *orientation session* on September 25, 1998. This session, which included program administrators and members of the testing team, addressed the purpose of the equating study and overall procedures for test administration. A two-hour *follow up training session* on October 23, 1999 dealt with more specific testing procedures and information about the CELSA and BEST Literacy. The third training session, held on October 29, was a two-hour *scoring workshop* involving the BEST Literacy test. After this workshop, four of the testing team members were selected to serve as scorers for the BEST Literacy test. Inter-rater reliability was calculated for this sub-group prior to the beginning of data collection on Nov. 3. The inter-rater reliability figures of the four scorers were high, with a mean of .99. In all, the testing team members underwent six hours of orientation and training before they began to collect data in Phase 1.

After data collection began in early November, a follow-up meeting involving members of the testing team was held on November 12. At this meeting, the team discussed issues which had come up during the initial week of testing and reviewed the procedures for administering the CELSA and BEST Literacy tests. All testing team members were trained to follow exactly the administration procedures for each test (as described in the test manuals), along with a set of additional procedures determined by the research director (see Appendix E). Each time the tests were given, the testing team members performed proscribed roles, such as primary tester and proctor (see Appendix E). Every effort was made to minimize variation that could be the result of testing conditions.

*Training for Phase 2:* In January 1999 a more extensive training regimen was implemented for the testing team as they prepared to administer the two oral tests (NYSPT and BEST Core). For this phase, the testing team members went through a three-part process for a total of nine hours of training.

Part 1, on January 6, was an overview of the BEST Core and NYSPT, with guided practice in giving both tests. The testing team members viewed the NYSPT videotape, observed a simulation of the BEST Core being administered, and practiced giving the tests to partners. Part 2 of the training, on January 8, involved additional practice time, discussion of the scoring system for each test, and scoring of test scripts (delivered orally) for the purpose of establishing inter-rater reliability. During Part 3, each testing team member was scheduled to practice administering the tests to ESL students from classes in the area. The

project and research directors monitored this practice session and evaluated each tester with a checklist in order to provide immediate feedback. (The data from Part 3 of the training were not included in the study.) In most cases, one session was sufficient for the testing team member to gain familiarity with the tests and give them confidently.

During the training sessions for both Phase 1 and Phase 2, two-way discussion and communication were encouraged. Members of the testing team made valuable suggestions for administration of the tests and were active in asking for clarification on areas such as scoring. As much as possible, the research and project directors attempted to clarify expectations through follow-up memos sent to the testing team after the training process was complete. (See Appendix G.) When the testing team was expanded to include several new members, this group underwent the same three-part process in late January and early February.

#### *D. Statistical Analysis Procedures.*

All of the data from the study were entered into a SPSS computer database for analysis. At the end of the data collection period, correlation coefficients were calculated for each pair of tests using the Pearson product-moment correlation. The significance level for these coefficients was set at a level of  $p < .01$  for a one-tailed test of significance.

After the correlation coefficients had been established, regression was calculated in order to predict relationships between scores on the pairs of tests. Along with the line of regression, the standard error of estimate and confidence intervals were calculated.

To establish ranges of scores on the CELSA and BEST for each SPL, the already-established BEST ranges were used to assign an SPL level to each student in the sample. The sample size, mean, and standard deviation were calculated for the CELSA or NYSPT scores in each level. Finally, the Multiple Classification Scheme (MCS), a form of centour analysis, was used to establish cutoff points between each level based on the figures for mean and standard deviation. Once these cutoff points had been set, they became the basis for the range of scores corresponding to each level presented in Tables 1 and 2.

The MCS is a statistical procedure for defining the cutoff point between two established categories which have overlapping scores. The formula for the MCS involves multiplying the mean of category A by the standard deviation of category B, adding this value to the mean of category B multiplied by the standard deviation of category A, and then dividing this value by the sum of both standard deviations (see Cooley and Lohnes, 1971; Ilyin, 1994). The resulting cut-off score is based on the intersection between the centours (or distribution of scores around the mean) of the two categories.

Using an alternative approach, the research director also calculated the formula for the regression line to determine predicted scores between the CELSA and BEST Literacy and between the NYSPT and the BEST Core. It is possible to use these predicted scores to establish ranges based on the previously-established cut-off points for the two versions of the BEST. However, it was felt that the MCS would provide a more accurate set of score ranges, especially at the lowest and highest levels.

The MCS was also used for equating the CELSA with SPLs 8 and 9. A correlation coefficient was calculated for the teacher ratings of students in this part of the study and their CELSA scores. Once a significant positive correlation was established, the students were grouped according to SPL 8 or 9. The mean and standard deviation were calculated for each group, and the MCS was used to establish cutoff points between SPLs 7, 8, and 9. Ranges of scores were set, completing the equating for the entire range

of CELSA scores and SPLs. The results of this statistical analysis will be presented in Section III.

#### ***E. Qualitative Analysis Procedures.***

Over the six-month data collection period, the testing team members gained valuable insights into each of the four tests used in this study. In order to collect and analyze these insights, the research and project directors used two qualitative procedures: (1) they created a survey which was distributed among the testing team members in March, and (2) they conducted a focus group interview with the members of the testing team in late April.

The survey collected information about the perceived strengths and weaknesses of each test as well as recommendations for use of the tests with specific programs or student populations (see Appendix L). As a follow-up to the survey, the research director developed a series of questions which were posed to the testing team members in a one-hour focus group interview on April 27. This interview was videotaped and transcribed.

The survey results and interview transcript were analyzed through a process of unitizing and coding the data through techniques common in qualitative research (Miles & Huberman 1994). This content analysis was used as one basis for Section 4 of this report, "Recommendations for Use of the Tests in Adult ESL Programs."

#### ***F. Limitations of the Study.***

The data collection procedures for this study were originally designed to have as little negative impact on the three participating agencies as possible. As a result, the testing conducted during Phase 1 was scheduled for outside of normal class times. Due to problems related to scheduling, the administrative team decided to conduct the testing during class times in Phase 2 and Phase 3. Despite this slight modification of the data collection procedures, the sampling procedures and the testing conditions remained consistent throughout the three phases of the study.

As the demographic figures show, the population tested with the CELSA during Phase 3 was different from the population tested in Phases 1 and 2. Students enrolled in the high level academic-professional classes (Advanced Grammar and Advanced Writing) tended to be more highly-educated, younger, more ethnically diverse, and more likely to be female than the students in the literacy ESL classes tested during Phase 2. However, one can make a case that the only students in the three agencies who would be rated at the 8-10 SPLs would be in the academic-professional program, since the descriptions for these levels involve proficiency in technical subject areas and because SPL 10 is, by definition, an educated native speaker of English.

Similarly, it must be noted that the research design was somewhat different in Phase 3 as compared with Phases 1 and 2. Instead of a random sample, the entire population of the highest-level academic-professional classes was tested. Instead of using BEST scores to determine the SPLs of the students, teacher ratings were used. These variations in the design were due to necessity. The population was small enough to be tested directly, and the BEST score ranges only covered SPLs 0-7.

Finally, it must be noted that the test results in this study were generated from a large, diverse sample of students in three suburban Chicago agencies in 1998-99. The population which was sampled may be somewhat different from the immigrant populations used in developing the original BEST/SPL equating

during the MELT project in the early 1980s. The population may also be somewhat different from the populations of agencies which might use the results of the study. For instance, the scales which resulted from this study may be more valid for a large, diverse community college program with a significant percentage of Hispanic students than for a small, homogeneous program with Southeast Asian students.

### III. Results and Discussion.

#### A. Supporting Data for the Results.

The score ranges for the CELSA and NYSPT in Table 1 were derived through a two-part process of (1) correlation/regression analysis to demonstrate a strong relationship between the test scores and (2) use of the Multiple Classification Scheme (MCS) to establish cutoff points between SPL levels. The statistical results which support this process are provided in the following section.

*Correlation and Regression.* As Table 5 demonstrates, the CELSA and BEST Literacy scores in this study had a strong positive level of correlation ( $r = .78$ ) which was found to be significant at the .01 level. The  $r^2$  value of .61 indicates that 61 % of the variance in the CELSA scores can be explained by the BEST Literacy scores. The strong positive correlation between the two tests is acceptable for using the BEST Literacy as the basis for equating the CELSA with the SPLs.

**Table 5: Statistical Analysis of BEST Literacy and CELSA**

TEST	N	Mean	SD	r	r <sup>2</sup>	Std. Error of Estimate	B	95 % Conf. Intervals
BEST-L	277	53.69	12.45	.78	.61	8.07	-12.32	-16.55 to -8.08
CELSA	277	30.81	12.84					

(Dependent variable: CELSA; Independent variable: BEST Literacy)

As Table 6 indicates, the relationship between the NYSPT and the BEST Core tests was even stronger, with a correlation coefficient of ( $r = .85$ ). This is a high enough correlation to directly equate scores one test with scores on the other test. The  $r^2$  value of .72 indicates the magnitude of the relationship between the two tests: at least 72% of the variation in the NYSPT can be explained by the BEST Core.

**Table 6: Statistical Analysis of BEST Core and NYS Place Tests**

TEST	N	Mean	SD	r	r <sup>2</sup>	Std. Error of Estimate	B	95 % Conf. Intervals
BEST-C	316	32.53	18.82	.85	.72	6.03	-5.078	-6.41 to -3.74
NYSPT	316	11.55	11.35					

(Dependent variable: NYSPT; Independent variable: BEST Core)



For further information about the correlation and regression, consult Appendix A for scatterplots and regression lines of the two score distributions.

*Data Used in MCS Analysis.* With a positive correlation established between the two pairs of tests, the BEST Literacy and Core tests were used to assign SPLs to each student in the sample. The mean and standard deviation for the CELSA and NYSPT were computed for each of these levels. Using the MCS formula, cut points were created between each SPL, and score ranges were established for each level. The data used in these calculations are provided in Table 7.

**Table 7 Descriptive Statistics for the CELSA and NYSPT at the SPLs**

SPL	CELSA N	CELSA mean	CELSA Std. Dev.	CELSA range	NYSPT N	NYSPT mean	NYSPT Std. Dev.	NYSPT range
0					36	1.03	1.38	
1					30	2.70	2.79	0-3
2	7	10.4	3.24	0-12	75	4.31	2.84	4-6
3	14	16.50	8.83	13-19	69	10.07	6.42	7-14
4	50	20.94	7.69	20-23	45	17.56	7.31	15-20
5	60	25.52	6.59	24-29	26	23.42	9.69	21-27
6	98	34.54	9.23	30-41	17	28.94	6.57	28-32
7	48	47.25	9.23	42-53	18	34.61	6.85	33+
8	60	59.28	9.99	54-64				
9	30	66.10	5.81	65+				

Given these results, a caveat is in order. Since the CELSA is a 75 item multiple-choice test, the lowest score ranges (for SPL 2 and 3) are questionable in terms of validity, because chance or guessing could achieve scores in this range. As a result, we recommend that the CELSA not be used for placement with students below SPL 4. This variation at the lowest level is less problematic for an oral test such as the NYSPT, where guessing is not much of an option. Consequently, we believe that the small score ranges for the NYSPT at Levels 1 and 2 do have some validity in placing students and reporting progress, although other factors (such as scores on other tests and student backgrounds) should also be taken into account by program administrators when making placements.

*Comparison with Other Language Proficiency Scales.* Beyond the high positive correlation between the pairs of tests, what supports the validity of these score ranges in relation to the SPLs? Although neither the NYSPT and the CELSA has been equated to the MELT SPLs prior to this study, both tests have been equated to English proficiency scales similar to the SPLs.



The four Student Performance Levels used in New York State have been equated with ranges of scores on the NYSPT. An examination of the descriptors for each level in this scale shows that they have much in common with the MELT SPLs. The score ranges for each SPL parallel closely the ranges identified in this study. (For this comparison, see the first table in Appendix C.)

Similarly, as the second table in Appendix C shows, the CELSA has been previously equated to a seven-level scale used in San Francisco, with scoring ranges that parallel the MELT SPLs in this study. The similarities in these test score ranges reinforce the validity of the results of this study.

### ***B. Demographic Profiles of the Three Agencies.***

As we have seen, the demographic profile each of the three agencies is somewhat different from the others. Since the results of this equating study relate to the aggregate of the three programs, we might ask to what degree the results reflect the individual character of each program. For a comparison of the three literacy programs and the academic-professional program at College of DuPage, see Table 8.

**Table 8: Demographic Comparison of the Three Agencies Including the Academic-Professional Classes at College of DuPage**

Agency	Sample size (N)	Languages in sample	Countries in sample	Mean age of sample	Mean years of education	Gender of sample (percent)
West Chicago D94	124	4	4	28.17	9.43	M 62.9 F 37.1
World Relief	146	17	23	34.99	9.99	M 41.8 F 58.2
College of DuPage Literacy Program	323	29	39	32.43	10.88	M 42.7 F 57.3
College of DuPage Acad.-Professional (Adv. Grammar and Writing classes)	103	27	37	28.99	13.62	M 27.2 F 72.8

The variations in the data among the three grant-funded adult ESL programs point to important differences in the character of the three agencies. The mean age of the World Relief students (approximately 35 years) is the highest of the three programs, indicating that the refugee population served by this agency includes a number of older students. In contrast, the program offered by West Chicago District 94 serves a younger (mean age 28), predominately male student population. The mix of students in the programs at World Relief and College of DuPage (with 23 and 27 countries represented, respectively) is far more heterogeneous than the predominately Hispanic student population in West Chicago District 94. A comparison of the most common languages and countries of origin in the sample from each program further clarifies this observation (see Table 9).

The educational background of the students in each program also varied somewhat. While each program included a broad range of students in terms of years of education, the lowest mean (9.43) was found at West Chicago District 94, while the highest mean (10.88) was identified at College of DuPage, with an overall mean of 10.35 years of education for the three programs. It should be noted that the sample did include a number of students with low levels of education. More than a quarter of students at West Chicago District 94 (27.4%) had six or fewer years of education, and nearly one-fifth (19.2 %) of the World Relief students fell into this category. The World Relief sample also included four students who had no prior education. (See Appendix B for a summary of these results.)

Table 9 compares the most common national backgrounds among the three programs. From the percentages in this table, we can see the homogeneous nature of the West Chicago District 94 program, and the overall predominance of ESL students from Mexico in DuPage County. In addition to immigrant groups such as Mexicans, World Relief's ESL program serves a refugee population which includes students from the Balkans (Bosnia, Yugoslavia, Croatia, Albania), Vietnam, and Sudan. The diverse population of the College of DuPage literacy program is reflected in 39 different countries of origin, including students from common Chicago-area immigrant groups (Mexico, Poland), students from East and Southeast Asia (China, Vietnam, Korea), and a broad range of non-Mexican Hispanics (Colombia, El Salvador, Guatemala).

**Table 9: Most Frequent Countries of Origin by Agency (percentages)**

West Chicago D94		World Relief DuPage		College of DuPage Literacy Program		Overall	
Mexico	97.6	Mexico	30.1	Mexico	52.6	Mexico	56.5
Colombia	0.8	Bosnia	19.9	Poland	8.7	Vietnam	6.1
Ukraine	0.8	Vietnam	13.7	Albania	5.3	Bosnia	5.6
Afghanistan	0.8	Yugoslavia	8.9	Vietnam	5.0	Poland	4.7
		Sudan	4.8	China	3.4	Albania	3.2
		Croatia	3.4	Afghanistan	3.1	Yugoslavia	2.7
		Colombia	3.4	Ukraine	3.1	Colombia	2.0
		Somalia	2.7	Korea	2.2	China	1.9
		Albania	1.4	Colombia	1.9	Ukraine	1.9
		Russia	1.4	El Salvador	1.5	Korea	1.3
		Iran	1.4	India	1.5	Sudan	1.2

A similar table comparing language backgrounds can be found in Appendix A.

These demographic comparisons provide a basis for understanding the population of students tested in this study and for generalizing the results to other programs. For each of the programs, the correlation coefficients for each pair of tests were in the .70-.90 range, meaning that there was a strong or marked relationship between the tests (see Appendix A). Although the results of this study were generated from the complete sample from the three programs (except in the case of equating the CELSA to SPLs 8 and 9), program directors can generalize the results of this study to specific programs in Illinois which are similar to the programs described here.

#### IV. Recommendations for Use of the Tests.

By allowing additional testing instruments to be used in ESL programs for tracking student progress, the Illinois State Board of Education has created more options for individual ESL programs in the state. With the results of this equating study in mind, we would like to suggest ways these tests can be most helpful for adult ESL programs in addition to noting some of the limitations of CELSA, NYSPT and BEST Literacy.

*The BEST Literacy* is the easiest test to use at the beginning level in large programs, since it can be group administered and makes distinctions between pre-literate (SPL 0-1) and literate beginning-level students (SPL 2-3). It can be used through the intermediate level (SPL 6) in programs which concentrate on life skill competencies. For more academically-oriented programs, a switch might be made at SPL 4 (or later) from the BEST Literacy to the CELSA, which can test up through SPL 9. (A student at SPL 10 is equivalent to a native speaker of English and thus does not need ESL services.) While the BEST Literacy test can be administered with minimal training to large groups of students at a time, it does require trained scorers for reliable scoring, and it is also somewhat time consuming to score (about 5 minutes per exam.). Additionally, it is a relatively expensive test to give in terms of the materials, with the booklet for each student costing \$1.75.

*The CELSA* should only be given to literate students at or above SPL 4. As we have noted, the scores for SPLs 2-3 could be achieved through chance. Like the BEST Literacy, the CELSA can be group administered, and it can be conveniently scored with Scantron answer sheets. Since the State carries the license for approved ESL programs, the CELSA is less expensive to administer than the BEST Literacy. In addition, the CELSA takes less time than the BEST, both to give and to score. However, the CELSA has its own pitfalls. The multiple choice cloze layout is difficult for some students to follow and, as with any multiple choice test, guessing is a possibility. Despite these problems, the CELSA does seem to be a good companion test to the BEST Literacy, and the high correlation between the two tests implies that a program could consistently record student progress if the BEST Literacy is used at the beginning and the CELSA at the intermediate levels.

*The NYSPT* does seem to make distinctions among students at SPL 0-3, although the scoring ranges are much smaller than those of the BEST at the lowest levels. (One might note that in New York's program, SPLs 0-3 are all subsumed into one beginning level; see Appendix C for a comparison.) The BEST Core and BEST Literacy do a more effective job of showing progress in the development of life skills, the usual focus for levels 0-3. However, the NYSPT is less expensive than the BEST, and it can be used effectively with small programs. Regardless of the size of the program, systematic training of a testing team is necessary to administer the NYSPT reliably.

The most important problem with the NYSPT is that it was developed as a *placement* test rather than as an ongoing measure of *progress* in an ESL program. Concerns about the NYSPT as a way of reporting ESL student progress were expressed in the qualitative data from the surveys and focus group interviews. Several ESL teachers and program administrators noted that the limited number of test items (27) and the nature of the pictures on this test did not reflect the life skill content of what is taught in a typical adult ESL literacy program. The NYSPT was perceived as having low content validity when it is used to evaluate student progress or program effectiveness.

Because the NYSPT's simple, reliable scoring system focuses on accuracy instead of fluency, the results of this test may be misleading as students in adult ESL programs increase in their spoken fluency. Longer, more fluent answers to the questions are more likely to contain errors and thus be recorded as a "1" instead

of a “2”. Some teachers in the focus group noted that it was possible for scores on the NYSPT to actually go *down* for individual students as they became more fluent in English.

A final objection to the NYSPT is that only one form of the test has ever been produced. To measure and report student progress, the same form (Form B) would have to be given repeatedly. Some students are likely to remember the pictures and questions from previous test-taking sessions, thus influencing the reliability of the results.

Although the NYSPT provides useful information about the oral proficiency level of students in adult ESL programs, our recommendation is that it only be used for placement purposes. The BEST Literacy and CELSA, when used together, provide a better picture of student progress over time from the pre-literacy to the advanced levels. This is, however, a necessarily limited picture, because the tests primarily focus on literacy and do not provide information about progress in oral communication.

## V. Conclusion.

This research project provides ESL teachers and program administrators with information which will help them interpret the scores of the CELSA and NYSPT tests and report the results of testing to funding agencies. In using these tests and the SPLs to report student progress, however, several points need to be kept in mind.

- ◆ The SPLs reflect levels of proficiency for both oral and written language. As a result, any student may be classified according to more than one SPL. For instance, it is possible for a student to take the NYSPT and score at SPL 2 and take the CELSA and score at SPL 4. *The NYSPT and the CELSA/BEST Literacy provide two very different sets of information and should not be used interchangeably.* For a more complete picture of a student’s English proficiency, a program may wish to give *both* the NYSPT and the BEST Literacy or CELSA.
- ◆ While the BEST Literacy and the CELSA correlate well and both tests focus on written language, these tests also have different emphases in content. The BEST Literacy tests mastery of life skills competencies related to reading and writing, while the CELSA assesses overall reading competence, including knowledge of a broad range of vocabulary and grammar. The BEST Literacy is most appropriate for determining the SPLs for students at the beginning level (SPLs 0-4), while the CELSA is most appropriately used with SPLs 5-7. Above SPL 7, most adult ESL programs are likely to use a test intended for native English speakers (such as the Test of Adult Basic Education, or TABE), although CELSA scores can be used to distinguish between SPLs 8 and 9.
- ◆ In conclusion, we recommend that programs use the BEST Literacy for reporting student progress in terms of SPLs at the beginning level, and the CELSA for reporting student progress at the intermediate level. These tests, used in conjunction, will at least show progress in terms of English literacy. We do *not* recommend using the NYSPT as a test of progress in English proficiency, although this test may be useful for placement purposes. Although the NYSPT correlates well with the BEST Core and can be equated with the oral SPLs, the limitations of this test make it problematic for reporting adult ESL student progress or evaluating the efficacy of adult ESL programs.

Finally, we would like to note that although the MELT SPLs are used widely, they are limited in their ability to distinguish between different levels of English proficiency, particularly in the area of literacy (*Research Agenda*, 1998. p. 10). We suggest that the Illinois State Board of Education (ISBE) and the Illinois Community College Board (ICCB) examine other forms of performance assessment, such as portfolio assessment, in order to evaluate the progress of adult ESL students.

## References

- Basic English Skills Test: Test Manual*. (1984). Washington, D.C.: Center for Applied Linguistics.
- Cooley, W.W. & Lohnes, P.R. (1971). *Multivariate Data Analysis*. New York: Wiley.
- Garreton, R. & Terdy, D. (1991). *Correlation Study of Adult English as a Second Language (ESL) and Adult Basic Education (ABE) Reading Tests*. Des Plaines, IL: Adult Learning Resource Center.
- Grognet, Allene, (1997). *Performance-Based Curricula and Outcomes: The Mainstream Language Training Project (MELT) Updated for the 1990s and Beyond*. English Language Training Technical Assistant Project. Denver: Spring Institute for International Studies.
- Ilyin, D. (1994). *User's Guide for CELSA*. Monecito, CA: Association of Classroom Teacher Testers.
- Miles, M. B. & Huberman, A.M. (1994). *Qualitative Data Analysis*, 2nd ed. Thousand Oaks: Sage
- The NYS Place Test: Administrator's Manual*. (1992). Albany: The University of the State of New York.
- Petersen, N.S. et al. (1989). "Scaling, Norming, and Equating." In *Educational Measurement*, 3rd e. Ed. R. Linn. New York: MacMillan (Chapter 6).
- Research Agenda for Adult ESL*. (1998). Washington, D.C.: Center for Applied Linguistics.

# Appendices

Appendix A	Statistical Information for Equating
Appendix B	Statistical Profiles: Program Demographics
Appendix C	Comparison of Score Ranges and Three Different Proficiency Scales
Appendix D	Training Schedule and Materials
Appendix E	Testing Team Instructions
Appendix F	Testing Schedule
Appendix G	Letters; to Teachers, Students
Appendix H	Consent to Test Forms
Appendix I	ISBD/TO's
Appendix J	SPL Checklist for Phase Three
Appendix K	Team Members
Appendix L	Recommendations for Testing in Various Programs

# Appendix A

## Statistical Information for Equating



ISBE Equating Study  
1998-1999

**Administrative Team:**

Joanna Escobar, College of DuPage  
Don Zabelin, West Chicago District 94  
Marilyn Sweeny, World Relief DuPage

**Research Director:**

Alan Seaman, Wheaton College  
**Project Director:**  
Linda Hayward, World Relief DuPage

**1. Purpose of the study.**

To equate the New York State Placement Test (NYS Place) and the CELSA with the Student Performance Levels (SPLs) for beginning and intermediate students (SPLs 0-7). The CELSA will also be equated with SPLs 8-10. *Using these results, program administrators in Illinois will be able to determine the SPL level of each adult ESL student who has taken the NYS Place, the CELSA, and the BEST Literacy tests.*

The data for this test equating study will be drawn from three DuPage County adult ESL programs: the College of DuPage, West Chicago District 94, and World Relief DuPage. These programs represent a broad range of adult ESL students and are similar to many other ESL programs in the state.

**2. Sampling.**

Between November and April, a sample of 250-300 intermediate-level ESL students from the three programs will take the CELSA and the BEST Literacy tests. During a second phase, between January and April, a sample of 250-300 beginning and intermediate level students will take the NYS Place and BEST Core tests. Finally, in April and May, a sample of 150 advanced academic-professional ESL students will take the CELSA and will be rated by their teachers according to descriptors created from SPLs 7-10.

**3. Test Administration**

To equate the scores of one test to another, it is imperative that the tests (NYS Place/BEST Core and CELSA/BEST Literacy) have a strong positive correlation. This can be achieved by comparing standardized tests which are similar in nature and by administering the test in a systematic and very consistent manner.

All of the tests will be administered by a highly-trained testing team under the supervision of the Project Director. The members of this team will undergo rigorous training and will have to demonstrate a high level of inter-rater reliability before they begin to collect the test data.

**4. Analysis**

The test scores will be analyzed statistically to establish a positive correlation between the tests and to equate test scores between the NYS Place and the BEST Core and the CELSA and BEST Literacy. These ranges of scores will be equated with the SPLs. The information on the tests will be reported for the entire sample and for subgroups within the sample, including ethnic identification, gender, length of time in the U.S., length of English study, and number of years of education.

The results from this study will be of particular interest to ESL program administrators in Illinois. In addition to equating the CELSA and NYS Place with the SPLs, the final report will contain recommendations for the appropriate use of each test in adult ESL programs, as well as recommendations for training teachers to administer the tests.

## Correlation and Regression

### *Tables, Scatterplots, Regression Lines*

**Descriptive Statistics**

	Mean	Std. Deviation	N
CELSA	30.81	12.84	277
BESTLIT	53.69	12.45	277

**Correlations**

	CELSA	BESTLIT
Pearson Correlation	CELSA 1.000	.779
	BESTLIT .779	1.000
Sig. (1-tailed)	CELSA .000	
	BESTLIT .000	
N	CELSA 277	277
	BESTLIT 277	277

**Descriptive Statistics**

	Mean	Std. Deviation	N
NYSPT	11.55	11.35	316
BESTCORE	32.53	18.82	316

**Correlations**

	NYSPT	BESTCORE
Pearson Correlation	NYSPT 1.000	.848
	BESTCORE .848	1.000
Sig. (1-tailed)	NYSPT .000	
	BESTCORE .000	
N	NYSPT 316	316
	BESTCORE 316	316

**Model Summary<sup>a,b</sup>**

Model	Variables		R	R Square	Adjusted R Square	Std. Error of the Estimate
	Entered	Removed				
1	BESTLIT <sup>c,d</sup>	.	.779	.606	.605	8.07

a. Dependent Variable: CELSA

b. Method: Enter

c. Independent Variables: (Constant), BESTLIT

d. All requested variables entered.

**Model Summary<sup>a,b</sup>**

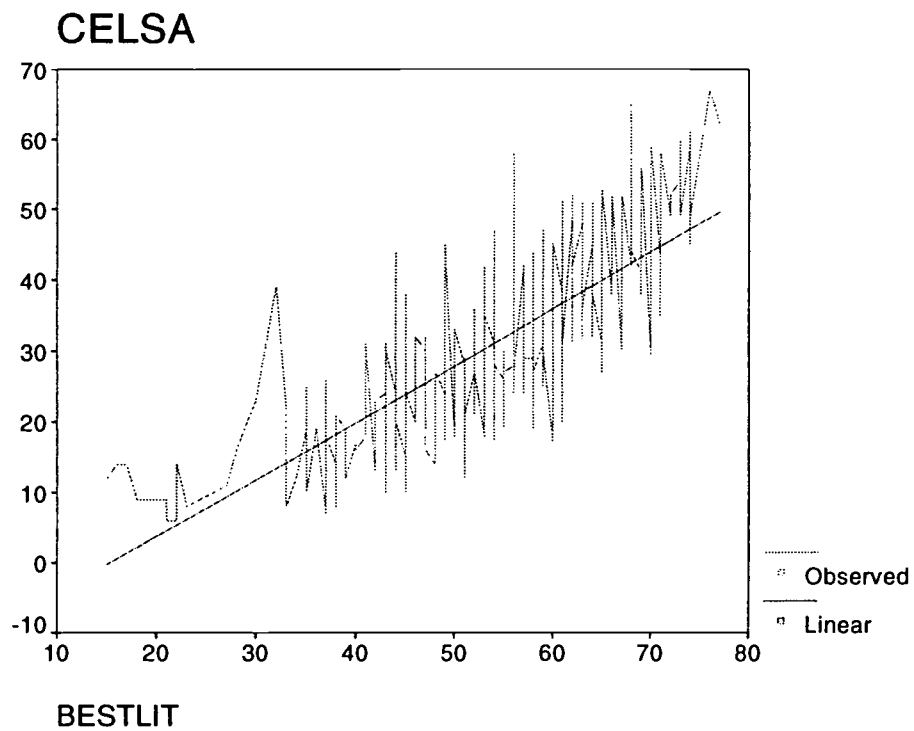
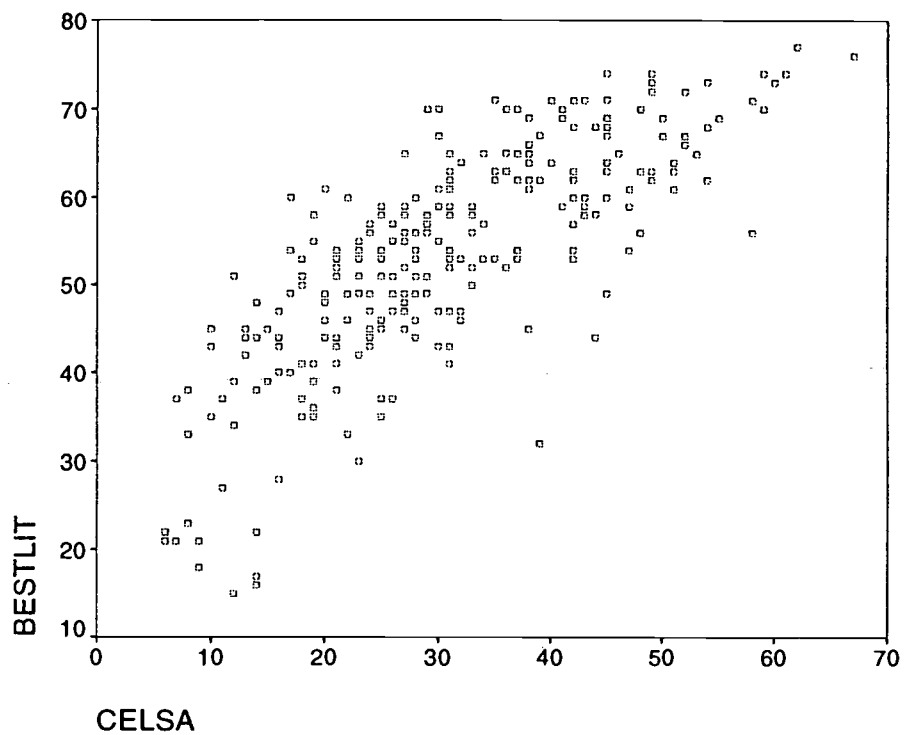
Model	Variables		R	R Square	Adjusted R Square	Std. Error of the Estimate
	Entered	Removed				
1	BESTCOR <sup>c,d</sup>	.	.848	.718	.718	6.03

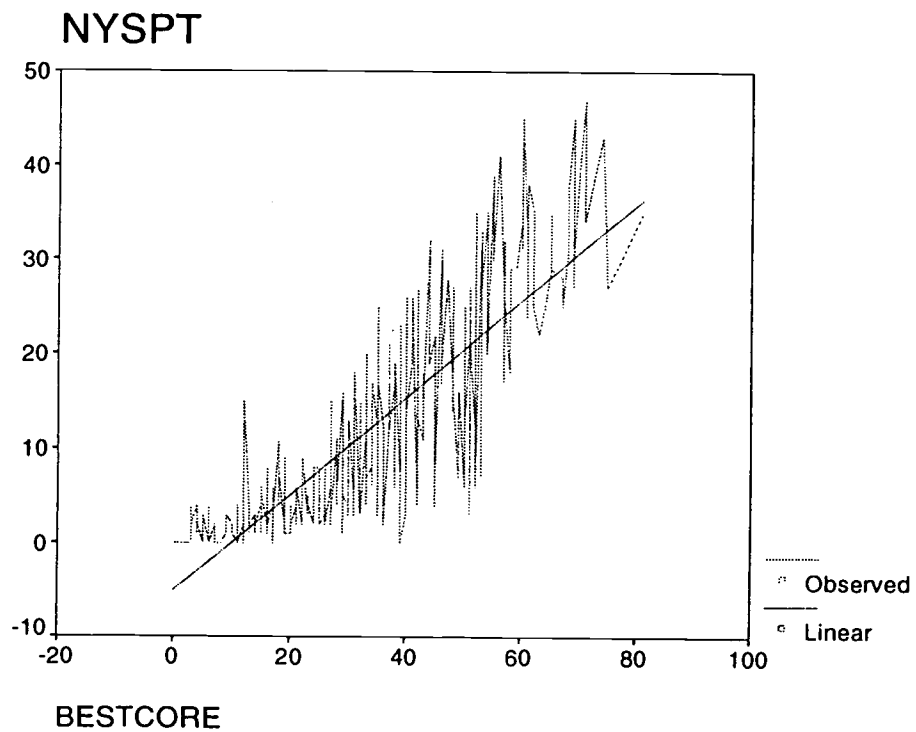
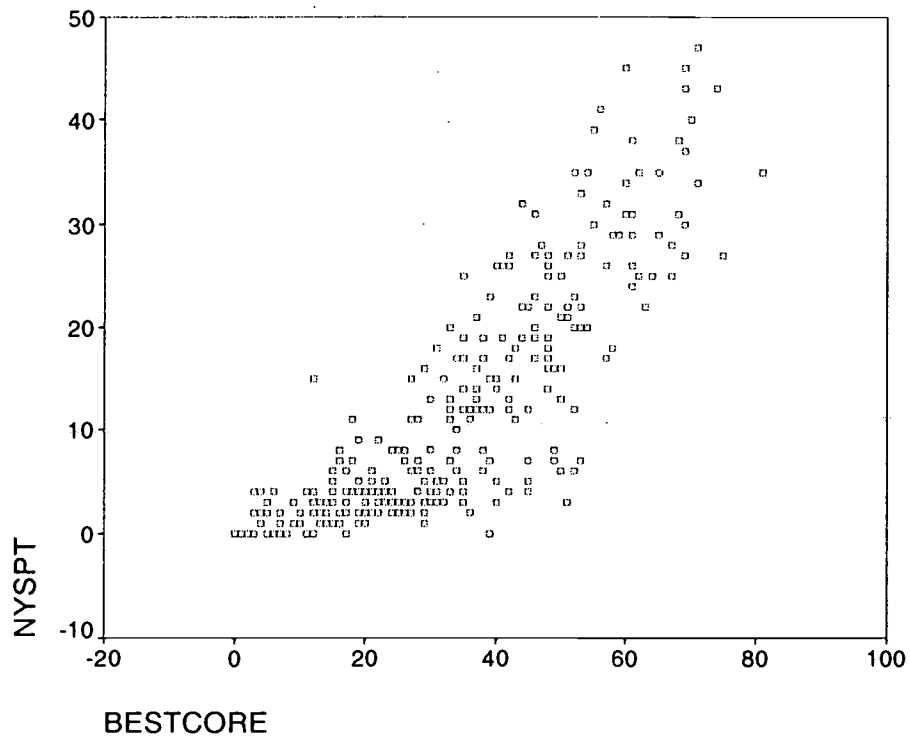
a. Dependent Variable: NYSPT

b. Method: Enter

c. Independent Variables: (Constant), BESTCORE

d. All requested variables entered.





## *Correlations Between the Pairs of Tests for Each Program*

### World Relief DuPage

**Correlations**

		BESTCORE	NYSPT
Pearson Correlation	BESTCORE	1.000	.890**
	NYSPT	.890**	1.000
Sig. (1-tailed)	BESTCORE	.	.000
	NYSPT	.000	.
N	BESTCORE	76	76
	NYSPT	76	76

\*\* . Correlation is significant at the 0.01 level (1-tailed).

**Correlations**

		BESTLIT	CELSA
Pearson Correlation	BESTLIT	1.000	.711**
	CELSA	.711**	1.000
Sig. (1-tailed)	BESTLIT	.	.000
	CELSA	.000	.
N	BESTLIT	70	70
	CELSA	70	70

\*\* . Correlation is significant at the 0.01 level (1-tailed).

### West Chicago Community High School District 94

**Correlations**

		BESTCORE	NYSPT
Pearson Correlation	BESTCORE	1.000	.849**
	NYSPT	.849**	1.000
Sig. (1-tailed)	BESTCORE	.	.000
	NYSPT	.000	.
N	BESTCORE	77	77
	NYSPT	77	77

\*\* . Correlation is significant at the 0.01 level (1-tailed).

**Correlations**

		BESTLIT	CELSA
Pearson Correlation	BESTLIT	1.000	.857**
	CELSA	.857**	1.000
Sig. (1-tailed)	BESTLIT	.	.000
	CELSA	.000	.
N	BESTLIT	47	47
	CELSA	47	47

\*\* . Correlation is significant at the 0.01 level (1-tailed).

### College of DuPage Literacy Program

**Correlations**

		BESTCORE	NYSPT
Pearson Correlation	BESTCORE	1.000	.829**
	NYSPT	.829**	1.000
Sig. (1-tailed)	BESTCORE	.	.000
	NYSPT	.000	.
N	BESTCORE	163	163
	NYSPT	163	163

\*\* . Correlation is significant at the 0.01 level (1-tailed).

**Correlations**

		BESTLIT	CELSA
Pearson Correlation	BESTLIT	1.000	.752**
	CELSA	.752**	1.000
Sig. (1-tailed)	BESTLIT	.	.000
	CELSA	.000	.
N	BESTLIT	160	160
	CELSA	160	160

\*\* . Correlation is significant at the 0.01 level (1-tailed).

<b>Inter-Rater Reliability</b>		
<b>SCORER</b>	<b>BEST Lit</b>	<b>NYSPT and BEST Core</b>
A	.992	.912
B	.999	.806
C	.997	.848
D	.996	.809
E		.797
F		.826
G		.802
H		.826
I		.853
J		.855
K		.882
L		.884
M		.832
N		.812
O		.788
P		.864
<b>TOTAL (MEANS)</b>	<b>.996</b>	<b>.837</b>

**Inter-rater reliability was calculated by (1) correlating the rater's score on each test item with a criterion score (Pearson r), and (2) averaging these correlations.**

## Appendix B

### Statistical Profiles: Program Demographics



## Demographic Information for the Three Programs

### Demographic Comparisons Among the Programs

PROGRAM	N	Education (years) mean/median		Time in US (years) mean/median		Length of English study (years) mean/median	
World Relief DuPage	146	9.99	11.00	2.77	1.00	1.05	0.50
West Chicago D94	124	9.43	9.00	7.82	3.00	4.10	0.75
College of DuPage	323	10.88	12.00	3.51	2.00	1.32	0.50
TOTALS (Literacy)	593	10.35	11.00	3.54	2.00	1.26	0.58

## Demographic Data: World Relief DuPage

		Country			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Mexico	44	30.1	30.1	30.1
	Uruguay	1	.7	.7	30.8
	Yugoslavia	13	8.9	8.9	39.7
	Iran	2	1.4	1.4	41.1
	Vietnam	20	13.7	13.7	54.8
	Eretria	1	.7	.7	55.5
	Bosnia	29	19.9	19.9	75.3
	Somalia	4	2.7	2.7	78.1
	Croatia	5	3.4	3.4	81.5
	Cuba	1	.7	.7	82.2
	Colombia	5	3.4	3.4	85.6
	Ethiopia	1	.7	.7	86.3
	Afghanistan	1	.7	.7	87.0
	Albania	2	1.4	1.4	88.4
	Korea	1	.7	.7	89.0
	Russia	2	1.4	1.4	90.4
	Ecuador	1	.7	.7	91.1
	Sudan	7	4.8	4.8	95.9
	Thailand	1	.7	.7	96.6
	Iraq	1	.7	.7	97.3
	Cambodia	1	.7	.7	97.9
	Germany	1	.7	.7	98.6
	Kazakhstan	2	1.4	1.4	100.0
	Total	146	100.0	100.0	
Total		146	100.0		

BEST COPY AVAILABLE

### Language

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Spanish	52	35.6	35.6	35.6
	Vietnamese	20	13.7	13.7	49.3
	Serbo-Cro	31	21.2	21.2	70.5
	Persian	2	1.4	1.4	71.9
	Tegma	1	.7	.7	72.6
	Bosnian	15	10.3	10.3	82.9
	Somali	3	2.1	2.1	84.9
	Amheric	1	.7	.7	85.6
	Russian	5	3.4	3.4	89.0
	Afghani	1	.7	.7	89.7
	Albanian	2	1.4	1.4	91.1
	Korean	1	.7	.7	91.8
	Arabic	7	4.8	4.8	96.6
	Thai	1	.7	.7	97.3
	Dinka	2	1.4	1.4	98.6
	Cambodian	1	.7	.7	99.3
	German	1	.7	.7	100.0
	Total	146	100.0	100.0	
Total		146	100.0		

### Education

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	4	2.7	2.7	2.7
	2	2	1.4	1.4	4.1
	3	4	2.7	2.7	6.8
	4	2	1.4	1.4	8.2
	5	5	3.4	3.4	11.6
	6	11	7.5	7.5	19.2
	7	6	4.1	4.1	23.3
	8	8	5.5	5.5	28.8
	9	17	11.6	11.6	40.4
	10	4	2.7	2.7	43.2
	11	12	8.2	8.2	51.4
	12	51	34.9	34.9	86.3
	13	4	2.7	2.7	89.0
	14	5	3.4	3.4	92.5
	15	3	2.1	2.1	94.5
	16	5	3.4	3.4	97.9
	17	2	1.4	1.4	99.3
	19	1	.7	.7	100.0
	Total	146	100.0	100.0	
Total		146	100.0		

### How Long US

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0-2m	12	8.2	8.2	8.2
	3-6m	34	23.3	23.3	31.5
	7-9m	16	11.0	11.0	42.5
	10-11m	7	4.8	4.8	47.3
	1y	18	12.3	12.3	59.6
	2y	13	8.9	8.9	68.5
	3y	11	7.5	7.5	76.0
	4y	6	4.1	4.1	80.1
	5y	7	4.8	4.8	84.9
	6y	5	3.4	3.4	88.4
	7y	2	1.4	1.4	89.7
	8y	3	2.1	2.1	91.8
	9y	3	2.1	2.1	93.8
	10y	1	.7	.7	94.5
	11-14y	3	2.1	2.1	96.6
	15-19y	4	2.7	2.7	99.3
	20+y	1	.7	.7	100.0
	Total	146	100.0	100.0	
Total		146	100.0		

### How Long ESL

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0-2m	27	18.5	18.5	18.5
	3-6m	52	35.6	35.6	54.1
	7-9m	22	15.1	15.1	69.2
	10-11m	4	2.7	2.7	71.9
	1y	16	11.0	11.0	82.9
	2y	14	9.6	9.6	92.5
	3y	1	.7	.7	93.2
	4y	4	2.7	2.7	95.9
	5y	1	.7	.7	96.6
	7y	1	.7	.7	97.3
	8y	2	1.4	1.4	98.6
	10y	1	.7	.7	99.3
	11-14y	1	.7	.7	100.0
	Total	146	100.0	100.0	
Total		146	100.0		

## Demographic Data: West Chicago Community High School District 94

### Country

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Mexico	121	97.6	97.6	97.6
	Colombia	1	.8	.8	98.4
	Ukraine	1	.8	.8	99.2
	Afghanistan	1	.8	.8	100.0
	Total	124	100.0	100.0	
Total		124	100.0		

### Language

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Spanish	119	96.0	96.0	96.0
	Russian	1	.8	.8	96.8
	Afghani	1	.8	.8	97.6
	Zopotezo	1	.8	.8	98.4
	Sapoteco	2	1.6	1.6	100.0
	Total	124	100.0	100.0	
Total		124	100.0		

### Education

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	2	1.6	1.6	1.6
	3	4	3.2	3.2	4.8
	4	1	.8	.8	5.6
	5	2	1.6	1.6	7.3
	6	25	20.2	20.2	27.4
	7	2	1.6	1.6	29.0
	8	8	6.5	6.5	35.5
	9	32	25.8	25.8	61.3
	10	3	2.4	2.4	63.7
	11	6	4.8	4.8	68.5
	12	19	15.3	15.3	83.9
	13	1	.8	.8	84.7
	14	4	3.2	3.2	87.9
	15	9	7.3	7.3	95.2
	16	5	4.0	4.0	99.2
	17	1	.8	.8	100.0
	Total	124	100.0	100.0	
Total		124	100.0		

### How Long US

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0-2m	1	.8	.8	.8
	3-6m	8	6.5	6.5	7.3
	7-9m	5	4.0	4.0	11.3
	1y	19	15.3	15.3	26.6
	2y	26	21.0	21.0	47.6
	3y	8	6.5	6.5	54.0
	4y	14	11.3	11.3	65.3
	5y	15	12.1	12.1	77.4
	6y	4	3.2	3.2	80.6
	7y	2	1.6	1.6	82.3
	8y	3	2.4	2.4	84.7
	9y	3	2.4	2.4	87.1
	10y	5	4.0	4.0	91.1
	11-14y	3	2.4	2.4	93.5
	15-19y	4	3.2	3.2	96.8
	20+y	4	3.2	3.2	100.0
	Total	124	100.0	100.0	
Total		124	100.0		

BEST COPY AVAILABLE



### How Long ESL

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0-2m	20	16.1	16.1	16.1
	3-6m	32	25.8	25.8	41.9
	7-9m	8	6.5	6.5	48.4
	1y	21	16.9	16.9	65.3
	2y	24	19.4	19.4	84.7
	3y	13	10.5	10.5	95.2
	4y	1	.8	.8	96.0
	5y	2	1.6	1.6	97.6
	6y	2	1.6	1.6	99.2
	11-14y	1	.8	.8	100.0
	Total	124	100.0	100.0	
Total		124	100.0		

## Demographic Data: College of DuPage

		Country			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Mexico	170	52.6	52.6	52.6
	Yugoslavia	3	.9	.9	53.6
	Vietnam	16	5.0	5.0	58.5
	Bosnia	4	1.2	1.2	59.8
	Somalia	1	.3	.3	60.1
	Colombia	6	1.9	1.9	61.9
	Ethiopia	1	.3	.3	62.2
	Ukraine	10	3.1	3.1	65.3
	Afghanistan	1	.3	.3	65.6
	Brazil	3	.9	.9	66.6
	Albania	17	5.3	5.3	71.8
	Poland	28	8.7	8.7	80.5
	Guatemala	3	.9	.9	81.4
	El Salvador	5	1.5	1.5	83.0
	China	11	3.4	3.4	86.4
	Korea	7	2.2	2.2	88.5
	Costa Rica	1	.3	.3	88.9
	Russia	3	.9	.9	89.8
	India	5	1.5	1.5	91.3
	Bangladesh	1	.3	.3	91.6
	Myanmar	1	.3	.3	92.0
	Taiwan	5	1.5	1.5	93.5
	Italy	1	.3	.3	93.8
	Japan	1	.3	.3	94.1
	Lithuania	3	.9	.9	95.0
	Venezuela	1	.3	.3	95.4
	Macedonia	1	.3	.3	95.7
	Czech R	2	.6	.6	96.3
	Ecuador	1	.3	.3	96.6
	Honduras	1	.3	.3	96.9
	Turkey	2	.6	.6	97.5
	Bulgaria	1	.3	.3	97.8
	Spain	1	.3	.3	98.1
	Syria	1	.3	.3	98.5
	Macedonia	1	.3	.3	98.8
	Lebanon	1	.3	.3	99.1
	Iraq	1	.3	.3	99.4
	Slovakia	1	.3	.3	99.7
	Armenia	1	.3	.3	100.0
	Total	323	100.0	100.0	
Total		323	100.0		

### Language

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Spanish	190	58.8	58.8	58.8
	Vietnamese	16	5.0	5.0	63.8
	Serbo-Cro	7	2.2	2.2	65.9
	Bosnian	1	.3	.3	66.3
	Somali	1	.3	.3	66.6
	Amheric	1	.3	.3	66.9
	Russian	6	1.9	1.9	68.7
	Afghani	1	.3	.3	69.0
	Portugese	3	.9	.9	70.0
	Albanian	17	5.3	5.3	75.2
	Ukranian	8	2.5	2.5	77.7
	Polish	27	8.4	8.4	86.1
	Chinese	16	5.0	5.0	91.0
	Korean	7	2.2	2.2	93.2
	Gujarati	2	.6	.6	93.8
	Bangla	1	.3	.3	94.1
	Burmese	1	.3	.3	94.4
	Italian	1	.3	.3	94.7
	Japanese	1	.3	.3	95.0
	Lithuanian	3	.9	.9	96.0
	Czech	2	.6	.6	96.6
	Turkish	2	.6	.6	97.2
	Malayalam	1	.3	.3	97.5
	Arabic	3	.9	.9	98.5
	Bulgarian	1	.3	.3	98.8
	Urdu	1	.3	.3	99.1
	Macedonian	1	.3	.3	99.4
	Slovensky	1	.3	.3	99.7
	Punjabi	1	.3	.3	100.0
	Total	323	100.0	100.0	
Total		323	100.0		

### Education

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	5	1.5	1.5	1.5
	1	3	.9	.9	2.5
	2	1	.3	.3	2.8
	3	1	.3	.3	3.1
	4	3	.9	.9	4.0
	5	7	2.2	2.2	6.2
	6	28	8.7	8.7	14.9
	7	8	2.5	2.5	17.3
	8	9	2.8	2.8	20.1
	9	48	14.9	14.9	35.0
	10	18	5.6	5.6	40.6
	11	26	8.0	8.0	48.6
	12	75	23.2	23.2	71.8
	13	19	5.9	5.9	77.7
	14	18	5.6	5.6	83.3
	15	19	5.9	5.9	89.2
	16	20	6.2	6.2	95.4
	17	10	3.1	3.1	98.5
	18	5	1.5	1.5	100.0
	Total	323	100.0	100.0	
Total		323	100.0		

### How Long US

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0-2m	7	2.2	2.2	2.2
	3-6m	46	14.2	14.2	16.4
	7-9m	27	8.4	8.4	24.8
	10-11m	10	3.1	3.1	27.9
	1y	55	17.0	17.0	44.9
	2y	44	13.6	13.6	58.5
	3y	40	12.4	12.4	70.9
	4y	25	7.7	7.7	78.6
	5y	9	2.8	2.8	81.4
	6y	8	2.5	2.5	83.9
	7y	5	1.5	1.5	85.4
	8y	8	2.5	2.5	87.9
	9y	6	1.9	1.9	89.8
	10y	8	2.5	2.5	92.3
	11-14y	11	3.4	3.4	95.7
	15-19y	10	3.1	3.1	98.8
	20+y	4	1.2	1.2	100.0
	Total	323	100.0	100.0	
Total		323	100.0		

### How Long ESL

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0-2m	76	23.5	23.5	23.5
	3-6m	87	26.9	26.9	50.5
	7-9m	16	5.0	5.0	55.4
	10-11m	4	1.2	1.2	56.7
	1y	64	19.8	19.8	76.5
	2y	32	9.9	9.9	86.4
	3y	13	4.0	4.0	90.4
	4y	5	1.5	1.5	92.0
	5y	7	2.2	2.2	94.1
	6y	9	2.8	2.8	96.9
	7y	1	.3	.3	97.2
	8y	3	.9	.9	98.1
	9y	2	.6	.6	98.8
	10y	1	.3	.3	99.1
	11-14y	3	.9	.9	100.0
	Total	323	100.0	100.0	
Total		323	100.0		

## **Appendix C**

### **Comparison of Score Ranges and Three Different Proficiency Scales**

**A Comparison of Score Ranges for the CELSA on Two Scales:  
The MELT SPLs and the Adult Education ESL Proficiency  
(The San Francisco Community College Centers)**

<b>CELSA Ranges</b>	<b>SPL Descriptions (Reading)</b>	<b>Adult Education ESL Proficiency Level Descriptions</b>	<b>CELSA Ranges</b>
0-19	<p><b>SPL 2</b> Functions in a very limited way in situations related to immediate needs. Recognizes letters of the alphabet, numbers 1-100, and a few very common sight words.</p> <p><b>SPL 3</b> Functions with some difficulty in situations related to immediate needs. Reads and understands a limited number of common sight words, and short, simple learned phrases related to immediate needs.</p>	<p><b>LOWER BEGINNING</b> - Designed for the student with little knowledge of English, the basic goal is for students to begin to communicate and conduct the affairs of daily life in spoken and written English. High frequency vocabulary is emphasized and regular forms of the present, past, and future are covered in the standard transformations.</p>	0-20
20-29	<p><b>SPL 4</b> Can satisfy basic survival needs and a few routine social demands. Reads and understands simple learned sentences and some new sentences related to immediate needs; frequent misinterpretations.</p> <p><b>SPL 5</b> Can satisfy basic survival needs and some limited social demands. Reads and understands some short, simplified materials related to basic needs with some misinterpretations.</p>	<p><b>UPPER BEGINNING</b> - Students continue to gain simple communicative ability to function in practical areas necessary for survival in the U.S. Although students read short authentic passages and simple conversations, emphasis continues to be on oral production and listening comprehension with an expansion of the structures of the lower beginning level.</p>	21-30
30-41	<p><b>SPL 6</b> Can satisfy most survival needs and limited social demands. Reads and understands simplified materials on familiar subjects. May attempt to read some non-simplified materials, but needs a great deal of assistance.</p>	<p><b>LOWER INTERMEDIATE</b> - A period of transition when students assimilate previously learned material and begin to develop fluency...Students read both short authentic materials as well as adapted reading selections.</p>	31-39

<b>CELSA Ranges</b>	<b>SPL Descriptions (Reading)</b>	<b>Adult Education ESL Proficiency Level Descriptions</b>	<b>CELSA Ranges</b>
42-53	<b>SPL 7</b> Can satisfy survival needs and routine work and social demands. Reads and partially understands non-simplified materials on everyday subjects; needs assistance.	<b>UPPER INTERMEDIATE -</b> Advanced grammatical structures such as complex sentences using embedding and relative clauses, passive forms, future perfect and continuous, unreal conditionals and complex word order are introduced. Students continue to read authentic materials and selections appearing in ESL textbooks.	40-48
54-64	<b>SPL 8</b> Can participate effectively in social and familiar work situations. Reads and understands most unsimplified materials, including materials in own field.	<b>LOWER ADVANCED -</b> Students learn to use more advanced structures... Written compositions and oral presentations are a basic part of the course... Students read authentic materials as well as text materials in ESL textbooks.	49-57
65+	<b>SPL 9</b> Can participate fluently and accurately in practical, social, and work situations. Reads non-simplified materials. Approximates a native speaker's ability to write accurately.	<b>UPPER ADVANCED -</b> Students review materials from previous courses and prepare for high school diploma courses, college entrance, or more advanced vocational training. A thorough review of English grammar, advanced reading skills, composition, and student research projects are basic components of this course.	58-66



**A Comparison of Score Ranges for the NYSPT on Two Scales:  
The MELT SPLs and the New York State ESOL Student Performance Levels  
(Source: NYS Place Test Administrator's Manual, pp. 11- 14)**

<b>NYSPT Ranges</b>	<b>SPL Descriptions (Listening and Oral Communication)</b>	<b>New York State ESOL Student Performance Level Descriptions</b>	<b>NYSPT Ranges</b>
0-14	<p><i>Level 0</i> - No ability whatsoever.</p> <p><i>Level 1</i> - Functions minimally, if at all, in English. Understands only a few isolated words, and extremely simple learned phrases. Vocabulary limited to a few isolated words. No control of grammar.</p> <p><i>Level 2</i> - Understands a limited number of very simple learned phrases, spoken slowly with frequent repetitions. Expresses a limited number of immediate survival needs using very simple learned phrases. Asks and responds to very simple learned questions. Some control of very basic grammar.</p> <p><i>Level 3</i> - Understands simple learned phrases, spoken slowly with frequent repetitions. Expresses immediate survival needs using simple learned phrases. Asks and responds to simple learned questions. Some control of very basic grammar.</p>	<p><i>Level 1</i></p> <p><i>Listening:</i> No ability whatsoever. Unable to understand spoken English except for a few isolated words and extremely simple previously-learned phrases. Able to understand a very restricted range of simple, previously-learned phrases spoken extremely slowly with frequent repetitions in familiar situations. Able to understand simple, previously-learned phrases spoken slowly with frequent repetitions in familiar situations.</p> <p><i>Oral Communication:</i> No ability whatsoever. Limited to a few isolated words. Adequate only to express a very restricted range of immediate needs using short, previously-learned phrases. Can respond to direct questions on familiar subjects using one or two words. Adequate only to express immediate survival needs using previously-learned, short phrases. Can respond to direct questions on familiar subjects using short, previously-learned phrases.</p> <p>Almost no control over basic grammar. Pronunciation frequently unintelligible.</p>	0-15

<b>NYSPT Ranges</b>	<b>SPL Descriptions (Listening and Oral Communication)</b>	<b>New York State ESL Proficiency Level Descriptions</b>	<b>NYSPT Ranges</b>
15-27	<p><i>Level 4</i> - Understands simple learned phrases easily, and some simple new phrases containing familiar vocabulary, spoken slowly with frequent repetitions. Expresses basic survival needs, including asking and responding to related questions, using both learned and a limited number of new phrases. Speaks with hesitation and frequent pauses. Some control of basic grammar.</p> <p><i>Level 5</i> - Understands learned phrases easily and short new phrases containing familiar vocabulary spoken slowly with repetition. Functions independently in most face-to-face basic survival situations but needs some help. Asks and responds to direct questions on familiar and some unfamiliar subjects. Still relies on learned phrases but also uses new phrases (i.e., speaks with some creativity) but with hesitation and pauses. Increasing, but inconsistent, control of basic grammar.</p>	<p><i>Level 2</i></p> <p><i>Listening</i> - Able to understand previously-learned phrases with ease and very simple new phrases which contain familiar vocabulary and are spoken slowly with frequent repetitions in familiar situations. Can partially understand new phrases spoken in contexts which help convey the meaning. Able to understand short phrases which contain familiar vocabulary and are spoken slowly, with repetition in both familiar and unfamiliar situations.</p> <p><i>Oral Communication</i> - Adequate to express basic survival needs using previously-learned phrases as well as some new phrases. Can ask and respond to direct questions on familiar subjects. Can engage in basic conversations on familiar subjects but lacks the ability to participate in most social situations. Speaks with obvious effort and frequent pauses.</p> <p>Very little control of basic grammar. Pronunciation often unintelligible.</p>	16-26
28-32	<p><i>Level 6</i> - Can satisfy most survival needs and limited social demands. Understands conversations containing some unfamiliar vocabulary on many everyday subjects, with a need for repetition, rewording, or slower speech. Functions independently in most survival situations, but needs some help. Participates with some confidence in social situations when addressed directly. Can sometimes clarify general meaning by rewording. Control of basic grammar evident, but inconsistent.</p>	<p><i>Level 3</i></p> <p><i>Listening</i> - Able to understand conversations on a variety of everyday subjects which contain some unfamiliar vocabulary and are spoken somewhat slowly with some need for repetition. Limited ability to function without face-to-face contact.</p> <p><i>Oral Communication</i> - Adequate to function in most face-to-face basic survival situations but needs help occasionally. Can ask and respond to direct questions on familiar subjects and a limited number of unfamiliar subjects. Can participate with difficulty in some social situations when addressed directly. Has limited ability to convey general meaning by repeating and re-wording but is usually unable to convey exact meaning or intentions. Shows some signs of spontaneity and creativity, speaks with obvious effort and frequent pauses. Increasing control of basic grammar evident but errors persist.</p>	27-34

# **Appendix D**

## **Training Schedule and Materials**

### Training Schedule:

#### Written Exams:

- September 25, 1998    Orientation session: The purpose of the project was presented.
- October 23, 1998    Detailed training for the CELSA and the BEST Literacy exams.
- October 29, 1998    Grading workshop, four members were selected to be the grading team, and their inter-rater reliability was calculated.
- November 12, 1998    Follow up meeting: discussion of initial testing, and review of the standardization procedures for the CELSA and BEST Literacy exams

#### Oral Exams:

- January 6, 1999 (January 22, 1999)    Overview of the orals; the NYSPT and the BEST Core, as well as initial hands on practice
- January 8, 1999 (January 29, 1999)    Additional practice time, clarification of grading procedures, collection of data for inter-rater reliability
- January 13, 14, 19, 1999 (February 3, 4, 8, 1999)    Administering the tests to real ESL students on a practice basis. Some testing team members were observed and rated with a qualitative checklist during each of these practice testing sessions, and in many cases were given immediate feedback as to positive observations or suggestions for change. They did not test students for data collection for the study until they had completed a practice session satisfactorily. In most cases, one session was sufficient for them to gain the familiarity and ease in administration to give the exams confidently and consistently.

There was a second round of oral training with additional members of the team that joined us from Wheaton College graduate program in TESOL in January. They followed the same procedures used for the first set of team members (dates in parenthesis).

Communication continued in both directions, with team members making suggestions or asking for clarification on procedures which were then communicated in follow up memos to the team.

## **ADMINISTRATION OF THE BEST AND NYS PLACE ORAL TESTS**

### **WORKSHOP, JAN 6, 2-5PM**

**I. Introduction, workshop and overview of tests and oral testing.**

Alan Seaman

**II. The NYSP**

Overview, walk through and video/Linda

Scoring/Alan and Linda

Practice/Alan

(Break, and refreshments)

**III. The BEST Oral**

Overview, walk through/Linda

Scoring/Alan and Linda

Practice/Alan

**IV. Question and Answer, on the Oral exams/Alan and Linda**

**IV. Brief review on duties, CELSA and BEST Lit administration/Alan and Linda**

**ORAL LANGUAGE TESTING PRACTICE SESSION 3**  
**NYS TEST/BEST ORAL**  
**Testing Team Comment Form**

Tester: \_\_\_\_\_

Date: \_\_\_\_\_

	too slow		natural pacing		too fast
I. TESTER'S SPEECH	1	2	3	4	5
					6
					7

	too serious/ intimidating		professional/ warm		overly friendly
II. TESTER'S MANNER	1	2	3	4	5
					6
					7

	overly rigid, mechanical		precise, natural		adding language
III. FOLLOWING TEST PROTOCOLS	1	2	3	4	5
					6
					7

Comments:

# Appendix E

---

## Testing Team Instructions

## TESTING TEAM JOB DESCRIPTIONS

### Beginning testing program:

**"Primary" testor:** This applies to each member giving the oral tests to the beginning proficiency students. Together we will work to set up the spaces for testing with the materials and to make the environment as conducive to the testing process as possible.

We will need to make sure that all the necessary materials are at each individual testing location before the testing begins. This will include the materials needed to administer the BEST Oral and the NYSP. For both exams, the testor will need pencils and a clipboard. For the BEST, the picture cue book, the testor cue book, the envelope with the correct money, a blank piece of paper to cover the cash and answer sheets are necessary. For the NYSP, the testing booklet, script, answer sheet, three pencils, and some small pieces of paper for each student to sign his/her name on are necessary. (These will all be brought to the test site by the Project Director. The testing team members need to what they need set up at their individual testing location.) Finally, each testor will need the ISBD's and translations to fill out with the student after giving the same student the second test. Since you will be the only one working with the student on the ISBD, it is critical that it be checked before dismissing the student to return to class. Please make sure that the address is complete (including the town and zip code) so the student can receive their stipend check. Additionally, the testor should check each piece of information to make sure the data is completely and accurately filled out. In particular, the birthdate needs to be checked, as some may write 98/99 for the year, or the test date rather than their birthdate. If a translation is available and is used, it makes errors less likely. Finally, the Primary Testor will fill out the TO (Testing Outcome) section of the ISBD's, and put the answer sheets together with the ISBD in a file for each student. A proctor may help with this after the testing is completed. These files need to be packed up to be returned to the PD's vehicle.

**Proctor:** This position will be quite different for this beginning testing. Again, this member will assist the with setting up and cleaning up the testing areas, and the refreshment area. The proctor will work as a scribe to document any irregularities observed during the testing process, and turn them in to the Project Director. The proctor will help with escorting students to their test location, and with the refreshments for students between the administration of the two oral exams. The proctor will also make sure the completed student files are delivered to the Project Director's vehicle.

Additional responsibilities to be shared by all testing team members at a site:

Assist in setting up and cleaning up the refreshment area for the students. This will be the "holding area" for the students participating in the study. They will be brought to this area until it is their turn to take their first test, and they will return to this area until they are given the second test. After they take their second test and fill out their ISBD, they may return to their classroom.

**Project Director:** This person is to oversee all aspects of the testing, and to help with all set up and clean up areas. The project director will also deliver and later collect in her car all testing and refreshment supplies, with the assistance of the other team members.

### Team members:

We will work cooperatively to cover the above responsibilities. Our goal as a team is to professionally and efficiently administer these tests to the students who are assisting us in this study. We want to do our job to the standards necessary for such a study, and to also appreciate the assistance we are receiving from the students that are enabling us to do so.



## **GUIDELINES FOR ADMINISTRATION OF WRITTEN ESL PROFICIENCY TESTS (CELSA, BEST Literacy, and others)**

- A. Be sure that the number of students taking the test fits the size of the room.** Students taking a test should not feel crowded or uncomfortable. Each student should be seated comfortably at a desk or table with plenty of room to write. Students should not be seated on the floor or in other unusual circumstances.
- B. Be sure that the conditions in the room are non-distracting.** Once the test begins, people should not be leaving or entering the room in large numbers. If a student needs to talk about something, have him/her come to the front of the room to consult quietly with a member of the testing team. If a member of the testing team converses with a student at his/her seat, the conversation should be very brief and whispered. No eating, drinking, or other distracting activity should take place during the test.
- C. Test instructions should carefully follow the guidelines in the testing manuals.** It is important that each "primary tester" use very similar wording in giving test instructions. Do not add your own elaborations or explanations of the test instructions. Do feel free, however, to answer any questions that the students might raise before the test begins.
- D. In giving test instructions, strive to be clear and also to put the students at ease.** It is important that the students feel comfortable as they begin each test, so do speak pleasantly to them while giving instructions. Remember to smile and thank them for coming. Speak naturally but slowly. If you are serving as the primary tester for a testing session, take time before the session to familiarize yourself with the test instructions. It's important that they are delivered naturally and clearly.
- E. Follow the allotted times for the tests very carefully.** When the time is up, collect the remaining tests from the students, even if they aren't finished. The maximum time allowed for each test should be the same. Let the students know when the tests will be collected. If possible, students should be able to see a clock while they are taking the test.
- F. Latecomers.** Begin the test instructions within 10 minutes of the scheduled time for the test. Students who come after the instructions are latecomers. You have two options for latecomers: (1) if there is a significant number of them, they can be seated in a separate room, given the same instructions by the primary tester, and allowed to take the test under supervision of a proctor; (2) if this isn't possible, or if they are more than 15 minutes late, they should be given information about an alternative time to take the test.
- G. The paperwork for each test (tests, answer sheets, student information forms) should be collected systematically.** If you are responsible for the answer sheets, you should make sure that each student's name is written clearly on each form, and that the forms are clipped together and given to the primary tester before everyone leaves.
- H. The tests must be scored systematically by designated scorers.** The CELSA answer sheets should be scanned by one or two designated people, with test sheets and results given promptly to the appropriate administrator. The BEST Literacy test forms should be scored by the designated scorers in one or two sittings for each group tested. (Do not score a few, do something else, come back later, score a few more, and so forth.) If you are scoring the BEST Literacy tests, set aside a block of time when you can concentrate on your work and score a group of tests. This should lead to more consistency in scoring.

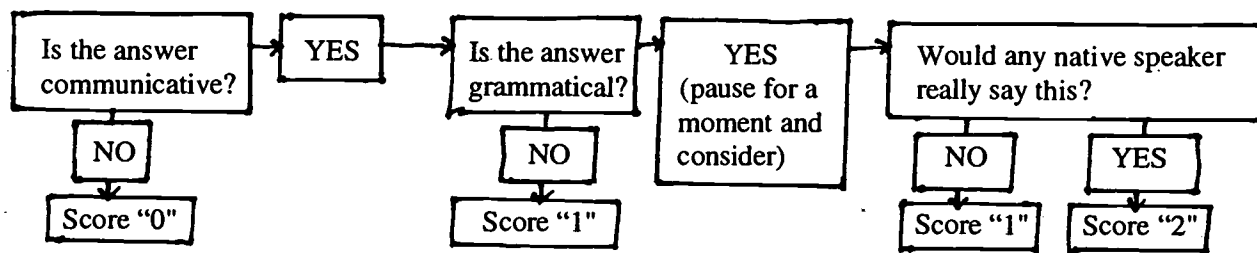
## GUIDELINES FOR ADMINISTRATION OF ORAL ESL PROFICIENCY TESTS (NYS Place, BEST Core, and others)

In addition to the test administration guidelines which are provided with each test, please keep the following points in mind.

1. **Consciously prepare for each test administration.** Go over the instructions and scoring for each test prior to giving it to the students. If a week or more has passed since you last gave the tests, look carefully at the test questions and practice saying them a bit. In other words, do whatever you can to be mentally prepared for the testing sessions. This will minimize variation in your own administration of the tests.
2. **Prior to giving the tests, set up the room in a manner which will minimize distractions.** Since we'll be testing in a wide variety of sites, this will involve some judgment calls. Make sure the students' backs are to each other, so they won't see other students being tested. If you are in a classroom with one or more other testers, sit as far as you can from the others to avoid distracting noise.
3. **Follow the testing script exactly.** Avoid giving any additional cues, such as non-verbal information or any "coaching." For the NYS Test, you can repeat the question once if the student doesn't understand.
4. **In your manner, be personable and try to put the students at ease.** Avoid an overly mechanical approach. Feel free to smile and provide affirming nonverbal communication. (Do this in moderation, of course.) During the third training session, we will be monitoring this aspect of your work on the testing team.
5. **Follow the standard protocols for the paperwork.** Check the scoring sheets and student information forms for accuracy before turning them in to the designated person. If the student is filling out the information form, please check it for accuracy. You may need to ask a follow-up question or two if the student didn't understand one or more parts of the form. It is also important that your name be on each of the scoring sheets that you turn in.

### SCORING THE NYS PLACE TEST

The most difficult aspect of scoring the NYS Place test involves distinguishing between a "1" and a "2" when the answer is grammatical. To consistently score these responses, testers should consider using the following flowchart. Begin by mentally registering a "2" for a grammatical response and then pause for a moment to consider whether or not any native speaker might give this response. If the response seems awkward or odd to you, score it as a "1."



*Alan A. Seaman, Ph.D. Wheaton Graduate School*

CELSA, Form 1  
Pre-test practice:

Directions to students:

Read the following quickly. Sometimes you see four words in a box. Choose the best word to complete the story or conversation. Read everything first. Don't write on the test paper. Write on the answer sheet. There are four answers for each question. Only one answer is correct. Fill in the letter of the correct answer. Here is an example:

— 11 A B C D E ● EXAMPLE:  
— 12 A B C D E

Practice Test

John is a student.

- X1. a. He  
b. He's  
c. She  
d. Him

studies English. He

- X2. a. is  
b. go  
c. likes  
d. happy

his

class. His teacher is very

- X3. a. work.  
b. school.  
c. good.  
d. day.

1 A B C D E  
2 A B C D E  
3 A B C D E

You will have 45 minutes to do the test. Work quickly. Do the easy questions first; then go back for the others. Do not use a dictionary or a book. After you finish, close the test. Stay in your seat at all times.

Any questions now - before you open the test.

OPEN THE TEST WHEN YOU HEAR THE WORD  
**BEGIN**

# Appendix F

## Testing Schedule

**Testing Schedule**  
**Equating Study, 1998-1999**

<b>Date(s)</b>	<b>Program (Number of classes)</b>	<b>Class Levels</b>	<b>Tests: Written, BEST Literacy and CELSA</b>	<b>Orals BEST Core and NYS</b>
<b><u>Fall:</u></b>				
11/3 and 5/98	West Chicago (4)	3-6	W	
11/9,10, 13/98	World Relief (4)	3-5	W	
11/16-19, 23/98	COD (10)	2-5	W	
11/30-12/2/98	COD (4)	multi, 5	W	
12/2-12/3	West Chicago (4)	3-6	W	
12/3,4/98	COD (1)	5	W	
12/7,9,11/98	COD (1)	3	W	
12/8,10,11,14,15/98	World Relief (5)	3-7	W	
<b><u>Winter:</u></b>				
1/20/99	COD (1)	1		O
1/20,25/99	West Chicago (4)	3-6	W	
1/25/99	COD (1)	3		O
1/26/99	COD (2)	1		O
1/28/99	COD (1)	1		O
2/2/99	COD (1)	2		O
2/3/99	COD (1)	multi		O
2/6/99	COD (1)	multi		O
2/8/99	World Relief (1)	1		O
2/9/99	COD (2)	1		O
2/11/99	COD (2)	2,3		O
2/16/99	COD (2)	1, multi		O
2/16,18/99	COD (2)	5, multi	W	
2/17/99	COD (2)	1,2		O
2/22/99	COD (2)	2,3		O
2/23,25/99	World Relief (2)	2,3	W	
2/24, 3/1/99	COD (3)	3,4	W	
2/25,3/1,2/99	COD (2)	4,5	W	
3/3,8/99	COD (3)	4,5,multi	W	
3/4/99	World Relief (2)	1		O
3/11/99	World Relief (1)	7		O
3/11,16/99	COD (2)	1,2	W	
3/15/99	West Chicago (4)	1,2		O
3/16/99	COD (2)	1,2		O
3/17/99	World Relief (1)	8		O
3/18/99	West Chicago (3)	1,2		O
3/22/99	World Relief (6)	1-6		O
	West Chicago (4)	1,3		O

Date	Program (#classes)	Level	Written	Orals
3/23/99	World Relief (3)	1,2,4		O
	West Chicago (2)	1,2		O
3/24/99	West Chicago (5)	1,2		O
<u>Spring:</u>				
4/6,8/99	World Relief (2)	2,3	W	
4/8/99	World Relief (6)	1,2,4-6		O
4/13/99	World Relief (1)	1		O
4/20/99	COD (2)	multi		O
4/22/99	COD (2)	multi		O
4/26,28/99	West Chicago (2)	2	W	
4/26/99	West Chicago (1)	4		O
4/28/99	West Chicago (1)	5		O
4/29/99	COD (1)	multi		O
5/4/99	COD (2)	advanced	CELSA	
5/12/99	COD (5)	advanced	CELSA	
5/13/99	COD (1)	advanced	CELSA	
5/18,20/99	COD (1)	1,2	W	

# Appendix G

## Letters; to Teachers, Students

## Equating Study Information Sheet

As ESL teachers, you are probably aware of the importance of standardized testing of your students, and it's relationship to funding. Illinois has recently approved two tests in addition to the CELSA; the New York State Placement Test (NYS Place) and the Basic English Skills Test (BEST/Literacy.) These tests will allow ESL programs to demonstrate gains for a larger range of our student populations, and thereby qualify districts for increased performance funding.

The West Suburban Area Planning Council is initiating a research project funded by a grant from the Illinois State Board of Education to equate the CELSA and the NYS Place with Student Performance Levels (SPL's.) Four hundred adult ESL students will be given the tests appropriate for their proficiency level; 200 from College of DuPage, 100 from West Chicago District #94, and 100 from World Relief DuPage. The students who participate must take their tests within one week in order to receive a \$25 stipend. Intermediate students will be given the CELSA and the BEST Literacy tests during November and December of 1998. Beginning level students will be tested from late January through mid-March, 1999. They will be given the NYS Place, and the BEST Oral tests on one day (both are one to one oral tests,) and will take the BEST Literacy test later during the same week.

The Research Director for this study is Dr. Alan Seaman of Wheaton College. He will be developing the criteria for a random but representative sampling of students from the programs. The Project Director is Linda Hayward, who will implement this study. She will work with Dr. Seaman, and the administrators of the three cooperating ESL Programs; Joanna Escobar at College of DuPage, Don Zabelin at West Chicago District #94, and Pam Meadows at World Relief DuPage.

If you would like additional information on this study, you may contact your program administrator, or Linda Hayward at World Relief DuPage, 462-7566.  
(email: [lhayward@ameritech.net](mailto:lhayward@ameritech.net))



Dear C.O.D. ESL teachers,

The Equating Study that we are participating in will soon be beginning the process of testing students. Before that process begins, we would like the students to be informed of the study. We have translated a student information letter into nine languages, with a simpler version for those whose languages were not interpreted. For those teaching on campus, we have requested that they would duplicate copies for their students in the appropriate languages. For those of you that do not have the facility to do that, would you please call me at World Relief DuPage and indicate which of the languages you need translated copies for, and how many of each. I will get them run off and on their way to you as soon as possible. My phone number is 462-7566, ext. 24. If I am out, please leave your name, class site, and the number you need for each language.

Some of your students will get a second letter inviting them to participate in this study. You will be asked to distribute, then collect and return those student responses to me as well. The invitations have been translated into the same languages. If you could indicate on a class list the languages you need and send that to me via Joanna Escobar, that would be very helpful when some of your students are identified to participate. There may be 20-30% of your class that will be invited to participate. We hope you will encourage those students that are chosen to help us with this study.

Sincerely yours,

A handwritten signature in cursive script that reads "Linda Hayward".

Linda Hayward, Project Director  
ISBE Equating Study

Illinois State Board of Education  
Equating Study Grant  
BEST/CELSA & BEST/NYS PLACE TESTS  
West Suburban Area Planning Council

College of DuPage, West Chicago District 94, Wheaton College, World Relief DuPage

A letter to ESL students at College of DuPage:

Our ESL program is helping to study ESL tests. We can have 100 beginning and 100 intermediate ESL students from our program in this study.

Each student that is chosen can earn \$25 for taking the tests. The beginning students will need to take two oral tests on one day. They will take an additional one hour test on another day later in the week. The intermediate level students will take two written tests. Each of these tests takes about one hour. The intermediate students will have to come two different times within one week.

Refreshments will be served. No childcare will be provided, so please don't bring any children with you. After you take the last test, the check for \$25 will be sent to your address.

The intermediate tests will be given in November and December 1998. The beginning tests will be given from January through March 1999. If you are chosen, you will need to come two times during one week in these months.

If you are chosen, you will get a letter. If you get a letter, we hope that you will be able to help us with this study. The letter will tell you where you need to come and what time the tests will begin.

If you have questions, please ask your teacher.

Thank you,

*Linda Hayward*

Linda Hayward  
Equating Study Project Director  
World Relief (630) 462-7566

વર્લ્ડ રીલીફ ઈએસએલ પિદાબાહિતો ૪૨નો ૪૩

અમારો અ.એસ.એલ. કાર્યક્રમ ઈએસએલ ૪૨ની મળવામાં મદદરૂપ  
બાબ છે. ઈએસએલ ના અભ્યાસ માટે અને આ કાર્યક્રમ માં ૫૦ પ્રાથમિક  
અને ૫૦ માધ્યમિક પિદાબાહિતો નો સમાવેશ કરીયો છે. ઈએસએલ  
ની ૪૨ની માટે નવ પિદાબાહિતો ની પસંદગી કરવામાં આવશે તેઓને બે  
મોનિટર ૪૨ની આપવી પડશે. તેઓ બીજી ટેકર ને એક ફીલ્ડ ની કોષ્ટક છે,  
તે કોઈ બીજા દિવસે પણ તેજ અકબાલીયા માં જોડા લેવાશે. માધ્યમિક  
અને ના પિદાબાહિતોને બે લેબીન ૪૨ની આપવી પડશે. આ દરેક ૪૨ની  
એક કલ્કુલેટ ની રૂંદો. માધ્યમિક પિદાબાહિતોને બે બુદા અને એક અકબાલીયા  
આપવાનું રૂંદો.

અભ્યાસી આપવામાં આવશે. બાપ્તિની વધાવદાર. લેવામાં જણી આવે. હવે  
૪૨ની આપવા ૪૨ની ૪૨ની રીલ નો એક નમારો આવનારો મોનિટર આપવામાં  
આવશે.

માધ્યમિક ૪૨ની આપવા નવેમ્બર, ડીસેમ્બર ૧૯૯૯ માં લેવાશે. શરૂઆતની  
૪૨ની આપવા નવેમ્બરની બી માસે ૧૯૯૯ બુદામાં લેવાશે. નો નમારો ૪૨ની  
કરવામાં આવશે નો નમારો આ નિદિગાઓમાં દર અકબાલીયા બે બાળ આપવાનું  
રૂંદો.

નમારો ૪૨ની નો નમારો ૪૨ની બુદામાં આવશે. નો નમારો કોઈ નમારો નો  
અને આપવા કે કે આ અભ્યાસ ને મદદરૂપ બધા માટે અકબાલીયા બધો. કોઈ  
બી નમારો કોઈ આપવાનું રૂંદો અને કોઈ ૪૨ની લેવાશે તે બુદામાં આવશે  
નમારો કોઈ પણ નમારો કોઈ નો મુદ્દાની કોઈ નમારો શ્રીમત ને પુરશે.

મુદ્દાની  
Jinda Hayward  
બોડા હુવડ

ઈસ્ટર્ન ૨૨ની પ્રોજેક્ટ કાયરેક્ટર  
વર્લ્ડ રીલીફ (૬૩૦) ૪૩૨-૭૪૬૬

BEST COPY AVAILABLE

Письмо студентам ESL в College of DuPage:

Наша программа помогает в исследовании тестов ESL. В этом исследовании могут принять участие 100 студентов ESL начального уровня и 100 – среднего.

Каждый из выбранных студентов получит \$25 за сдачу тестов. Студенты начального уровня должны будут сдать два устных теста в один день и ещё один тест в другой день на той же неделе. Студенты среднего уровня будут сдавать два письменных теста, каждый из которых длится около одного часа, в два разных дня на одной неделе.

Участвующим в тестах будут предложены напитки и закуски. Присмотр за детьми не предусмотрен, поэтому не берите с собой детей, пожалуйста. После сдачи последнего теста чек на \$25 будет выслан по вашему адресу.

Тесты среднего уровня будут проводиться в Ноябре-Декабре 1998 г. Тесты начинающего уровня будут проводиться с Января по Март 1999 г. Если Вы будете выбраны, то должны будете прийти дважды в одну из недель соответствующего месяца.

Если Вы будете выбраны, то получите письмо, содержащее информацию о месте и времени проведения тестов. Надеемся что Вы сможете помочь нам в этом исследовании.

Если у вас есть вопросы, обратитесь к Вашему учителю.

Спасибо,

Linda Hayward  
Equating Study Project Director  
World Relief (630) 462-7566

# Appendix H

## Consent to Test Forms

Please, come and take some ESL tests!  
We need your help.

You have been selected to be part of a very important study for the State of Illinois. We are studying ESL tests.

Between 100 and 150 students from World Relief will help. We would like you to be one of the students. You will take some special ESL tests during one week. If you take two tests, we will pay you \$25.

We would like you to take the tests at (site) \_\_\_\_\_

on (date)\_\_\_\_\_ at (time)\_\_\_\_\_

and on (date)\_\_\_\_\_ at (time/site)\_\_\_\_\_

Please mark one of the following boxes:

- ☐ Yes, I will help you with this study.
- ☐ No, I'm sorry. I cannot help you with this study.

Name: \_\_\_\_\_

Address: \_\_\_\_\_

\_\_\_\_\_

Phone: \_\_\_\_\_

Please return this form to your teacher this week!

**¡Venga y toma los exámenes de inglés!  
Necesitamos su ayuda.**

Ud. es escogido para ayudar en un estudio importante por el estado de Illinois.  
Estudiamos los exámenes de inglés como segunda lengua.

Solo 100 estudiantes de West Chicago District #94 ayudarán. Queremos que Ud. sea uno de estos estudiantes. Ud. tomará algunos exámenes especiales durante una semana. Si Ud. toma todos los exámenes, le pagaremos \$25.

Se servirán refrescos. No hará cuido de niños, y por eso es importante que Ud. no traiga niños.

Queremos que Ud. tome los exámenes en (lugar) \_\_\_\_\_ el (fecha) \_\_\_\_\_  
\_\_\_\_\_ a (tiempo) \_\_\_\_\_ y el (fecha) \_\_\_\_\_  
a (tiempo) \_\_\_\_\_

Favor de escoger uno de los siguientes:

\_\_\_\_\_ ¡Si! Yo puedo ayudarles con este estudio

\_\_\_\_\_ No, lo siento. No puedo venir en estas fechas o tiempos, pero  
podría venir:

(Escoge otra fecha y tiempo del horario incluido. Ambos tienen que ser en una semana)

(lugar) \_\_\_\_\_ (tiempos) \_\_\_\_\_

(fechas) \_\_\_\_\_

\_\_\_\_\_ No, lo siento. No puedo ayudarles con este estudio.

Nombre: \_\_\_\_\_

Dirección: \_\_\_\_\_  
\_\_\_\_\_

Teléfono: \_\_\_\_\_

**¡Entrega este papel a su maestra esta semana!**

## MOLIMO VAS DA UCESTVUJETE U ESL TESTIRANJU!

### POTREBNA NAM JE VASA POMOC

Vi ste odabrani da ucestvujete u vrlo vaznom istrazivanju za drzavu Illinois. U tom istrazivanju se proucavaju ESL testovi.

Samo 200 studenata sa World Relief ce ucestvovati. Mi bi zeleli da vi budete jedan od tih studenata. Za to bi trebalo da uradite neke ESL testove u toku jedne nedelje. Mi cemo vam platiti 25 dolara ako uzmete sve te testove.

Bice servirano neko manje posluzenje. Nije obezbedjeno cuvanje dece., tako da vas molimo da ne dovodite decu.

Mi bismo zeleli da uzmete testove (mesto) \_\_\_\_\_  
u (datum) \_\_\_\_\_ u (vreme) \_\_\_\_\_

Molimo vas da oznacite jednu od sledećih rubrika:

- ☐ Da! I mogu da dodjem I pomognem vam u vasem istrazivanju
- ☐ Ne, zao mi je. Ne mogu vam pomoci u to istrazivanju.

Name: \_\_\_\_\_

Adresa \_\_\_\_\_

Telefon: \_\_\_\_\_

Molimo vas da vratite ovaj formular vashoj uciteljici u toku ove nedelje!



**请来参加英语测试!**  
**我们需要您的帮助!**

您已被选中参加ILLINOIS州的一项重要研究项目: 英语作为第二语言测试研究.

**World Relief** 仅有200名学生参加此项研究, 希望您是其中的一名. 您将于一周内参加一些英语考试. 在全部考试完成后, 您将得到25元美金的报酬.

考试中心备有茶点, 但不设儿童活动中心, 届时请不要携带儿童.

考试地点 \_\_\_\_\_  
第一次考试日期 \_\_\_\_\_ 时间 \_\_\_\_\_

请选择:

☐ 我愿意参加此项测试研究.

☐ 对不起, 我无法参加此项测试研究.

姓名: \_\_\_\_\_  
地址: \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

电话: \_\_\_\_\_

**\*请将此表于本周交还给您的老师!**

# **Appendix I**

## **ISBD/TO's**

**Individual Student Bio-Data  
Testing Outcomes**

West Suburban Area Planning Council - Equating Study, ISBE Grant  
Intermediate Student Identification Form

**Name:** \_\_\_\_\_

**Address:** \_\_\_\_\_

**Program:** College of DuPage      West Chicago #94      World Relief

**Teacher:** \_\_\_\_\_ **Class level:** \_\_\_\_\_

**Sex:** Male \_\_\_\_\_ Female \_\_\_\_\_ **Birthdate:** Month \_\_\_\_\_ Day \_\_\_\_\_ Year 19 \_\_\_\_\_

**Country of origin:** \_\_\_\_\_ **First language:** \_\_\_\_\_

**Education:** (Number of years of schooling completed in your native country) \_\_\_\_\_

**How long have you been in the United States?** \_\_\_\_\_

**How long have you studied English?** \_\_\_\_\_

---

**Test scores:**

Test # 1: (Circle)      BEST      CELSA

Date: \_\_\_\_\_ 1998

Score: \_\_\_\_\_

Rater: \_\_\_\_\_

Comments of rater: \_\_\_\_\_

---

Test # 2: (Circle)      BEST      CELSA

Date: \_\_\_\_\_ 1998

Score: \_\_\_\_\_

Rater: \_\_\_\_\_

Comments of rater: \_\_\_\_\_

West Suburban Planning Council - Equating Study, ISBE Grant  
Beginning Student Identification Form:

**Name:** \_\_\_\_\_

**Address:** \_\_\_\_\_

**Program:** College of Dupage      West Chicago #94      World Relief

**Teacher:** \_\_\_\_\_ **Class level:** \_\_\_\_\_

**Sex:** Male \_\_\_\_\_ Female \_\_\_\_\_ **Birthdate:** Month \_\_\_\_\_ Day \_\_\_\_\_ Year 19 \_\_\_\_\_

**Country of origin:** \_\_\_\_\_ **First language:** \_\_\_\_\_

**Education:** (Number of years of schooling completed in your native country) \_\_\_\_\_

**How long have you been in the United States?** \_\_\_\_\_

**How long have you studied English?** \_\_\_\_\_

**Test # 1:** (Circle)      **NYS Place**      **BEST Oral**

**Date:** \_\_\_\_\_ 1999

**Score:** \_\_\_\_\_

**Rater:** \_\_\_\_\_

**Comments of rater:** \_\_\_\_\_

**Test # 2:**      **NYS Place**      **BEST Oral**

**Date:** \_\_\_\_\_ 1999

**Score:** \_\_\_\_\_

**Rater:** \_\_\_\_\_

**Comments of rater:** \_\_\_\_\_

West Suburban Planning Council - Equating Study, ISBE Grant  
Academic/Professional Student Identification Form

**Name:** \_\_\_\_\_

**Address:** \_\_\_\_\_

**Teacher:** \_\_\_\_\_ **Class:** \_\_\_\_\_

**Sex:** Male \_\_\_\_\_ Female \_\_\_\_\_ **Birthdate:** Month \_\_\_\_\_ Day \_\_\_\_\_ Year 19\_\_\_\_

**Country of Origin:** \_\_\_\_\_ **First language:** \_\_\_\_\_

**Education:** (Number of years of schooling completed) \_\_\_\_\_

**How long have you been in the United States?** \_\_\_\_\_

**How long have you studied English?** \_\_\_\_\_

**CELSA:**

**Date:** \_\_\_\_\_

**Score:** \_\_\_\_\_

**SPL:**

**Score (level):** \_\_\_\_\_

**Rater:** \_\_\_\_\_

**Comments:** \_\_\_\_\_

\_\_\_\_\_

# **Appendix J**

## **SPL Checklist for Phase Three**

## Questionnaire for Advanced Students in Academic-Professional Program ISBE Test Equating Study

### *English Language Proficiency Ratings*

Before answering these questions, please look over any records you have of this student's performance in your class. Concentrate for a half-minute or so on your memories of interaction with the student during the past four weeks.

Based on your experience with this student, which statements best describe his or her performance with the English language? Each language skill is classified in four categories (A through D). Read the descriptions of *all* of the categories, and then classify the student according to the descriptors which fit him or her best. Note that the categories begin at a high intermediate/low advanced level (Category A) and progress up to Category D, where the student's performance is equal to that of a native speaker of English. In other words, the categories here are the highest levels of an overall scale.

We have used **boldface** and *italics* to highlight the key words which involve comparisons among the various categories. Your categorization of the student should be based on a *general impression* of which set of descriptors *best* fits him or her. As this student's teacher, you are the person best qualified to make this impressionistic judgment.

Thank you for your time and effort on this important project. When you have completed the forms, please return them to the project director, Linda Hayward.

## GENERAL LANGUAGE ABILITY

### Category A

- ♦ The student can use English to satisfy survival needs and routine work and social demands.
- ♦ The student can handle work that involves following oral and written instructions in familiar and some unfamiliar situations.
- ♦ A native English speaker not used to dealing with limited English speakers can *generally* communicate with this student on familiar topics.

### Category B

- ♦ The student can participate *effectively* in social and familiar work situations.
- ♦ A native English speaker not used to dealing with limited English speakers can communicate with this student on *almost* all topics.

### Category C

- ♦ The student can participate *fluently and accurately* in practical, social, and work situations.
- ♦ A native English speaker not used to dealing with limited English speakers can communicate easily with this student.

### Category D

- ♦ The student's ability is equal to that of a native English speaker of the same socio-economic level.

## Questionnaire for Advanced Students in Academic-Professional Program ISBE Test Equating Study

Student Name: \_\_\_\_\_ Class: \_\_\_\_\_  
Teacher: \_\_\_\_\_ Date: \_\_\_\_\_

## LISTENING COMPREHENSION

### Category A

- ♦ The student understands conversations on most everyday subjects at normal speed when addressed directly; he or she may need repetitions, rewording, or slower speech.
- ♦ The student understands routine work-related conversations.
- ♦ The student has difficulty following conversation between native speakers.

### Category B

- ♦ The student understands general conversation and conversation on technical subjects in his or her own field.
- ♦ The student understands without face-to-face contact (telephone, TV, radio); he or she may have difficulty following rapid or colloquial speech.
- ♦ The student understands most conversation between native speakers; he or she may make details if speech is very rapid or colloquial or if subject is unfamiliar.

### Category C

- ♦ The student understands *almost* all speech in any context. He or she is occasionally confused by highly colloquial or regional speech.

### Category D

- ♦ The student's listening comprehension is equal to that of a native English speaker of the same socio-economic level.



## ORAL COMMUNICATION

### Category A

- ♦ The student functions independently in survival and many social and work situations, but may need help occasionally.
- ♦ The student communicates on the telephone on familiar subjects.
- ♦ The student expands on basic ideas in conversation, but still speaks with hesitation while searching for appropriate vocabulary and grammar.
- ♦ The student clarifies general meaning easily, and can sometimes convey exact meaning.
- ♦ The student controls basic grammar, but not more difficult grammar.

### Category B

- ♦ The student participates effectively in practical and social conversation and in technical discussions in his or her own field.
- ♦ The student speaks fluently in both familiar and unfamiliar situations; he or she can handle problem situations.
- ♦ The student conveys and explains the exact meaning of complex ideas.
- ♦ The student has good control of grammar.

### Category C

- ♦ The student approximates a native speaker's fluency and ability to convey ideas precisely, even in unfamiliar situations.
- ♦ The student speaks without effort.
- ♦ The student has excellent control of grammar with no apparent patterns of weakness.

### Category D

- ♦ The student's oral communication is equal to that of a native speaker of the same socio-economic level.

3

## READING

### Category A

- ♦ The student reads and partially understands some non-simplified materials on everyday subjects; he or she needs assistance.

### Category B

- ♦ The student reads and understands most non-simplified materials, including materials in his or her own field.

### Category C

- ♦ The student reads non-simplified materials.

### Category D

- ♦ The student's reading is equal to that of a native speaker of the same socio-economic level.

## WRITING

### Category A

- ♦ The student performs routine writing tasks within a familiar context.
- ♦ The student makes some errors; he or she may need assistance.

### Category B

- ♦ The student performs writing tasks with reasonable accuracy to meet social and basic work needs.

### Category C

- ♦ The student approximates a native speaker's ability to write accurately.

### Category D

- ♦ The student's writing is equal to that of a native English speaker of the same socio-economic level.

4

# Appendix K

## Team Members

## **ISBE Equating Study Personnel**

### **Research Director:**

Dr. Alan Seaman  
Director TESOL  
Wheaton College  
Wheaton, IL 60187

### **Project Director:**

Linda Hayward  
World Relief DuPage  
1825 College Ave, Suite 230  
Wheaton, IL 60187

### **Program Administrators, Members of the West Suburban Area Planning Council:**

Joanna Escobar  
Director of Adult Education  
College of DuPage  
22 Street and Lambert Road  
Glen Ellyn, IL 60137

Don Zabelin  
Director of Adult Education  
West Chicago District 94  
326 Joliet Street  
West Chicago, IL 60185-3142

Marilyn Sweeny  
Director  
World Relief DuPage

Pamela Meadows  
ESL Coordinator  
World Relief DuPage

### **Testing team members:**

Diane Brown  
Dorothy Carter  
Ginger Dakin  
Gail Davis  
Sonja Gassett  
Marilyn Gebhardt  
Allison Geijer  
Heather Hasty  
Lois Hill  
Christine Kim  
Jennifer Kline  
Antonio Lollino  
Kathy Marsh  
Nita Martindale  
Kathleen Poach  
Innes Sheridan  
Marion Schulte  
Dolores Ward

### **Assistant to the Project Director**

Esther Joung

## Appendix L

### Recommendations for Testing in Various Programs

## **The BEST Literacy, CELSA, and NYS Place Tests: OVERALL RECOMMENDATIONS FOR ADULT ESL PROGRAMS**

### **A. For Smaller Programs:**

- \* The *NYS Place* is easier to implement if a group of teachers has been trained to administer it simultaneously. It covers beginning to advanced levels.
- \* The *BEST Literacy* is also a viable option. Train several testers who will specialize in scoring.
- \* The *CELSA* has less of a complete range than the other two, but makes finer distinctions among students at the intermediate and advanced levels

### **B. For Large Programs:**

- \* The *NYS Place* may be too difficult to implement logistically unless a testing team is trained and funded.
- \* The *BEST Literacy* can also be time-consuming unless fairly large groups are tested at the same time
- \* The *CELSA*'s format may be difficult for some ESL students and it cannot be used effectively below SPL 3. However, it is the easiest to implement and score in large programs.

### **C. For Non-Formal and Non-Academically-Oriented Programs:**

- \* Does the *NYS Place* track progress well in non-academic programs? It doesn't contain much life skills information, although it will show gradually-increasing command of structures and oral communication skills.
- \* The *BEST Literacy* may be a good option if literacy/life skills are the focus. It may be problematic if there are many adult students who have been in the US a long time -- this can skew the results upward
- \* The *CELSA* may be the least relevant test for students in this category.

### **D. For Formal and Academic-Professional Programs:**

- \* The life-skills emphasis of the *BEST Literacy* may be seen as off-putting or irrelevant by these students, particularly beyond the beginning level
- \* The *NYS Place* and *CELSA* have more of a structural focus. They might be used in tandem at the intermediate and advanced levels to track student progress.



**U.S. Department of Education**  
Office of Educational Research and Improvement (OERI)  
National Library of Education (NLE)  
Educational Resources Information Center (ERIC)



## **NOTICE**

### **Reproduction Basis**



This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.



This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").